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

Ran Zhuo
Ross School of Business, University of Michigan, Ann Arbor, MI 48105, ranzhuo@umich.edu
Audrey Tiew
Department of Economics, New York University, New York, NY 10012, audreytiew@nyu.edu

High-tech manufacturing industries face constant technological change and sustained investment pressures as manufacturing technologies evolve. National policies from the U.S. and China—driven by concerns over security and industrial self-reliance—now heavily influence these investment pressures. This paper examines the effects of industrial policy on contract manufacturing investments in the global semiconductor industry. We assemble a novel dataset that combines quarterly facility-level capacity investments with global contract manufacturing orders from 2004 to 2015. Using these data, we estimate a structural model of contracting between semiconductor manufacturers and their clients, recovering key competitive parameters. We apply the model to a detailed case study of a major semiconductor manufacturer evaluating whether to locate a large fabrication facility in the U.S. or remain in its home region amid shifting industrial policies. In counterfactual simulations, we find that locating in the U.S. would require an additional investment of $1.2 billion compared to the home region, which could roughly be offset by lump-sum subsidies comparable (percentage-wise) to those provided under the CHIPS and Science Act. However, the profit reduction caused by cost disadvantages and reduced competitiveness at the U.S. location has a greater impact. Across various policy scenarios motivated by real-world industrial policies—including a no-policy baseline, U.S. export controls alone, Chinese tax subsidies alone, or a combination of both—locating in the U.S. consistently results in an additional $1.6–$1.8 billion profit loss compared to staying in the home region. This loss is concentrated in later years as the facility's technology matures. U.S. import tariffs substantially alter these patterns: under our conservative tariff rate calibration, relocating to the U.S. increases the firm's profit by $5.9 billion relative to remaining in its home region, primarily by offsetting the cost disadvantages of a U.S. facility and enhancing the firm's competitiveness in capturing U.S. demand, particularly for mature technologies.

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, which emphasize domestic self-reliance in key high-tech sectors. These policies aim to boost domestic production through generous manufacturing subsidies, tax incentives, export controls, and potentially 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 distorting manufacturers' incentives for capacity investment across regions. While industry experts generally welcome subsidies, they caution that broad restrictions intended to decouple the global supply chains of these highly internationalized industries could undermine their long-term health and growth (see, e.g., Semiconductor Industry Association 2022, 2021).

In this paper, we examine how industrial policies in the U.S. and China may influence semiconductor manufacturers' regional capacity investment decisions. Using realistic policy bundles from both countries, we ask: If a particular set of policies had been in place during our data period, how should a manufacturer assess facility location choices to maximize long-run profit? While numerous reports from government agencies, industry groups, think tanks, and market intelligence firms have analyzed the implications of recent U.S. and Chinese policies for the semiconductor supply chain—and offered recommendations to both national governments and firms—there remains a surprising lack of data-driven, evidence-based academic research on this topic.

Most policy-oriented reports are highly qualitative, offering conclusions and recommendations based on limited macroeconomic or sector-level data (e.g., Friedberg and Boustany Jr 2020, Semiconductor Industry Association 2021, Bateman 2022, U.S. Government Accountability Office 2022, to name a few). The most data-driven analyses include Kim and VerWey (2019), which examines the potential impact of China achieving chip self-reliance; Varas and Varadarajan (2020) and U.S. Chamber of Commerce (2021), which study the effects of trade policies such as export controls and tariffs; and Varas et al. (2020), which analyzes the impact of manufacturing subsidies. In particular, 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.

Among the few academic studies, Jacobs et al. (2022) examine the stock market impact of the 2018 U.S. ban prohibiting American firms from supplying to ZTE, a major Chinese telecommunications company. Funke and Wende (2022) theoretically evaluate the macroeconomic effects of U.S. semiconductor export restrictions, while Park and Liu (2023) empirically investigates similar issues using multi-regional input-output data. Han et al. (2024) construct measures of technological decoupling and interdependence between the U.S. and China using patent data, and study the effects of U.S. export restrictions on the performance and innovation of Chinese firms in sanctioned supply chains.

Although the above reports and studies provide valuable insights, they typically focus on individual policies or narrow sets of interventions. Few, if any, consider how combinations of policies interact to shift market equilibrium in this highly globalized and intensely competitive industry. To analyze the effects of multiple industrial policies and their interactions and to understand firm decision-making in this complex setting, a model-based approach that explicitly captures competitive equilibrium is

essential. Two recent studies move in this direction. Goldberg et al. (2024) documents recent industrial policies in the global semiconductor sector and employs a model-based approach to detect unmeasured subsidies and to quantify how such subsidies can lower production costs through learning-by-doing. Miao (2024) develops and estimates a dynamic oligopoly model in which firms decide whether to innovate and adopt more advanced manufacturing technologies, explicitly incorporating trade disruption risk and industrial policy. To maintain tractability, the model abstracts from dynamic capacity investment, treating capacity as frictionlessly re-optimizable over time.

To the best of our knowledge, we are among the first to empirically examine how industrial policies aimed at promoting domestic self-reliance affect the capacity investment decisions of semiconductor manufacturers.[1] We address this important research gap by combining detailed industry data with a rigorous structural model of market competition among manufacturers. Although fragmented, the data landscape in the semiconductor industry is rich and highly granular. Academic research has drawn on this detail to study a wide range of topics—including productivity (e.g., Irwin and Klenow 1994, Salomon and Martin 2008), resource allocation (e.g., Chang and Matsumoto 2022), competition (e.g., Henderson 1993, Fuchs and Kirchain 2010), innovation (e.g., Hall and Ziedonis 2001, Eizenberg 2014), vertical relationships (e.g., Terwiesch et al. 2005, Kapoor 2013, Fontana and Greenstein 2021), and market power (e.g., Gugler and Siebert 2007)—across the various product markets within the semiconductor supply chain. From a methodological perspective, several studies have estimated structural models to analyze industry dynamics in the semiconductor and electronics sectors, focusing on competition and investment decisions (Goettler and Gordon 2011, Gardete 2016, Igami 2017, 2018, Igami and Uetake 2020, Thurk 2022, Igami et al. 2024). The markets examined in these papers share key features with our focal market, including oligopolistic competition among a few large firms, a unified global market, vertical product differentiation, and sustained, capital-intensive capacity investment driven by technological progress.

Our empirical setting focuses on the contract manufacturing market in the semiconductor industry. Contract manufacturers produce chips on behalf of other companies, primarily fabless designers that lack their own fabrication facilities. Contract manufacturers play a vital role in the global semiconductor supply chain, producing chips from microprocessors for laptops to system-on-chips for smartphones and AI accelerators for large language models. Their clients include leading chip designers like Apple and Nvidia as well as customers requiring legacy chips for defense and medical applications. The market is highly concentrated: in 2017, contract manufacturing generated over $62 billion in revenue, with Taiwan Semiconductor Manufacturing Company (TSMC) alone accounting for more than half of the top eight manufacturers' revenue. 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.

Contract semiconductor manufacturing has become a central focus of recent industrial policies in both the U.S. and China, which prioritize domestic capacity expansion through substantial direct subsidies. The U.S. CHIPS and Science Act of 2022, for example, allocates $39 billion to support

[1] Bollinger et al. (2024) studies a related question in the context of solar panels.

the construction and modernization of domestic fabs, alongside a 25% investment tax credit. TSMC alone has been awarded $6.6 billion in subsidies for its $65 billion investment in the multi-phase TSMC Arizona facility, which targets cutting-edge manufacturing. Meanwhile, China has supported its semiconductor industry since 2014 through the "Big Fund," which has invested around $50 billion, primarily in manufacturing capacity. Other policy tools employed or proposed by both countries in recent years include export controls that restrict manufacturers from contracting with Chinese buyers, as well as tax breaks and tariffs. The rapid pace and complexity of these policy changes make it difficult to isolate the effects of individual interventions using reduced-form methods like event studies. This further motivates the use of a structural, model-based approach that explicitly captures market equilibrium to analyze the combined effects of policies and firm strategies.

The data we analyze combines various rich data sources from the industry. It includes a comprehensive, worldwide sample of facility-level capacity investments, captured quarterly from 1995 to 2015. This sample provides details on facility characteristics, installed capacity, manufacturing technology, ongoing construction and equipment investments, and the type of investment (e.g., constructing a new facility or upgrading an existing line). Additionally, we have a representative sample of global contract semiconductor manufacturing orders, collected quarterly from 2004 to 2015. This sample includes data on order quantity, price, manufacturing facility location, and technology. We enhanced these two proprietary datasets with additional, publicly available information on industry and firm performance. This supplementary information covers annual revenue figures, revenue shares by manufacturing technology generation and geographic region, gross margins, R&D spending, and capacity utilization for top contract manufacturers.

We use these data to estimate a structural model of capacity contracting between manufacturers and their clients, recovering the key parameters that determine market equilibrium during the 2004–2015 period. Accordingly, our counterfactual analysis of facility location choices also focuses on this period. This timeframe precedes the announcement or implementation of most recent industrial policies, so the industry was more likely to be in a stable equilibrium to enable reliable inference. Although national semiconductor policies have changed frequently, the shifts prior to 2015 were relatively minor compared to the significant policy changes in China and the U.S. beginning in late 2014. Moreover, the fundamental economics of the industry have not changed significantly from this period to the present. We thus believe that analyzing and exploring counterfactual scenarios from this recent historical period not only remains interesting and important but also offers useful insights for policymakers and firms today.

Our model endogenizes manufacturer profits and how they vary with the number of competitors and capacity constraints in the contract semiconductor manufacturing industry through a stylized bargaining framework. In this framework, buyers solicit bids from multiple manufacturers with available capacity for a given technology generation and period. Buyers have inelastic demand for chip quantities, reflecting downstream rigidity, and choose manufacturers to maximize their utility based on willingness to pay, possible regional preferences or subsidies, and the price charged.

Manufacturers' price bids depend on both the competitive environment and their marginal costs. When only one manufacturer bids, it sets a take-it-or-leave-it price equal to the buyer's maximum willingness to pay adjusted for regional manufacturing utility. When multiple manufacturers bid, negotiation follows an English auction where the winner prices just above the second-best competitor's marginal cost, adjusted for the buyer's regional preferences. Each manufacturer's marginal costs include common technology costs, manufacturer-specific cost advantages, capacity scale effects, regional cost differences, and private idiosyncratic shocks. Total profits aggregate over all won contracts, accounting for price, costs, and fixed overhead. This framework effectively links competition, capacity constraints, and regional factors to pricing and profitability in contract semiconductor manufacturing.

The key competitive parameters recovered from estimation include buyers' regional preferences, manufacturer-specific marginal costs, regional differences in marginal costs, and cost advantages from geographic proximity between manufacturers and buyers. These parameters are estimated using the method of simulated moments. In each iteration, simulated buyer-manufacturer negotiations generate model-implied moments, which are compared to observed moments such as gross margins, price distributions, revenue, utilization, and regional revenue shares. The estimation minimizes a weighted distance between simulated and observed moments (rescaled to make metrics like revenue and utilization comparable) using weights that balance the contribution of moment groups with differing numbers of observations. This approach enables robust identification of structural parameters that align with observed variation in prices, quantities, and firm-level outcomes across time, technologies, and regions.

The estimation results confirm key economic intuitions and provide credible quantitative insights into the global contract semiconductor manufacturing market from 2004 to 2015. The estimated parameters show that producing more advanced technologies is more costly, while costs decline over time as technologies mature. Firm- and region-specific cost differences indicate that major players like TSMC enjoy substantial cost advantages—both from firm-level efficiency (with a marginal cost that is, on average, 13.96% lower than that of the average manufacturer in Taiwan) and from regional cost savings (with a marginal cost 9.43% lower than if the same fab were located in North America). The model also identifies a significant cost benefit—on average equivalent to 7.96% of marginal cost—when buyers contract with manufacturers in the same region, underscoring the role of geographic proximity in shaping market outcomes. On the demand side, the estimates show that North American buyers dominated the market for cutting-edge technologies. However, their share has steadily declined over time and is much smaller for mature technology generations, as mainland Chinese buyers have become increasingly prominent.

We then apply the model and the estimates to a detailed case study of TSMC evaluating whether to locate a large fabrication facility in the U.S. or remain in its home region amid shifting industrial policies. While ideally we would analyze TSMC's 2020 decision to build fabs in Arizona, the Arizona facilities were developed amid frequent policy changes, so the market likely operated outside a stable equilibrium needed to enable reliable inference of key cost and competitive parameters in the baseline. To overcome these challenges, we identify comparable projects within our sample period of 2004 – 2015. In particular, TSMC Fab 14 in Taiwan closely matches Arizona's investment profile in many

ways, including development timeline, phased strategy, and adoption of cutting-edge technology with preplanned upgrades. The main difference is that the Arizona fab is located outside TSMC's home region and benefits from substantial government subsidies. This makes Fab 14 our ideal sandbox for simulating counterfactual scenarios, comparing maintaining Fab 14 in Taiwan versus relocating Fab 14 to the U.S., and studying how industrial policy bundles—such as subsidies, tax credits, export controls, and tariffs—affect the optimality of location decisions and competitive outcomes. These results provide insight into how policies might influence investments like TSMC Arizona.

Our counterfactual analyses of TSMC Fab 14's location choices yield several insights. First, based on facility characteristics (such as capacity, location, and technology) and the construction and equipment costs of similar projects in our facility investment data, we find that locating in the U.S. would require an additional $1.2 billion investment compared to the home region. These additional costs could be roughly offset if lump-sum subsidies—comparable in percentage terms to those provided under the CHIPS and Science Act—were available.

However, the reduction in profit caused by cost disadvantages and reduced competitiveness at the U.S. location has a more significant impact. Across various policy scenarios inspired by real-world industrial policies—including a no-policy baseline, U.S. export controls alone, Chinese tax subsidies alone, and a combination of both—locating in the U.S. consistently results in a profit loss of $1.6 – $1.8 billion compared to remaining in the home region. This loss is concentrated in later years, as the facility's technology matures. In the early years, competition is less intense, with fewer manufacturers capable of producing frontier technologies; proximity to high-end U.S. customers helps offset the higher production costs. Over time, as low-end demand shifts to Asia and global competition intensifies—since more manufacturers can produce at mature technology nodes—U.S.-based production becomes increasingly uncompetitive. Across these realistic policy scenarios, we find that locating in the U.S. is not a profitable strategy.

Tariffs, however, substantially alter these patterns. When Fab 14 remains at home, tariffs have a muted effect on firm profits. Given limited U.S. capacity during the sample period and inelastic buyer demand, U.S. buyers have little choice but to contract with foreign-based fabs and absorb the additional tariff-related costs. In contrast, when Fab 14 is located in the U.S., tariffs help offset the cost disadvantages of the U.S. facility and enhance the firm's ability to competitively capture U.S. demand, particularly at the low end—demand it previously had to compete for globally against low-cost overseas manufacturers. Under our conservative tariff rate calibration, relocating to the U.S. increases the firm's profit by $5.9 billion compared to remaining in its home region. In this context, tariffs effectively counterbalance the competitive disadvantages typically faced by fabs producing mature technologies in high-cost locations like the U.S., resulting in the opposite pattern as compared to that observed under other policy combinations.

Our findings highlight the critical role of technology maturity, the differing competitive environments and market structures between advanced and mature technologies, and evolving buyer demographics in shaping high-tech manufacturing investment decisions amid intensifying geopolitical pressures. While the short-term impact on profitability may be limited when technology is at the cutting

edge, industrial policies are poised to exert a much stronger influence over the long term as technologies mature and market structures evolve. Moreover, across the various policy bundles we simulated and different potential facility locations, we consistently observe that industrial policies—such as export controls and tax incentives enacted by rival countries—adversely affect TSMC's profitability. Under our conservative calibration scenarios, these policies could reduce profits by approximately $3.1 billion to $10.6 billion. This significant impact underscores that the expanding scope and intensity of industrial policies are likely to fundamentally reshape firms' long-term profitability outlooks. Consequently, firms must critically reassess their capacity investment strategies, factoring in not only current market conditions but also the evolving geopolitical landscape and competitive environment as technology matures, to remain competitive and resilient.

2. Empirical Setting

This section starts with an introduction to contract manufacturing of semiconductor chips, describing where contract manufacturing situates in the industry supply chain and the major players. We then describe more precisely the product being manufactured by contract manufacturers, the manufacturing process, and the distinction between different generations of manufacturing technologies. This will help us understand the fundamental economics of the industry. We then discuss the recent policy developments in China and the U.S. regarding domestic manufacturing of semiconductor chips, covering the main policy instruments considered or implemented by national governments. We wrap up the section by introducing the case study that anchors our analysis: a major semiconductor manufacturer (TSMC) evaluating whether to locate a large fabrication facility in the U.S. or remain in its home region amid shifting industrial policies.

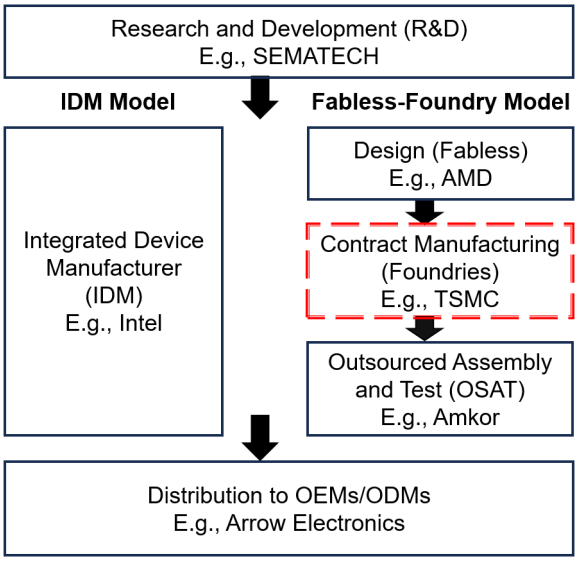
2.1. Contract Manufacturing in Semiconductor Industry Value Chain

The semiconductor industry has a complex value chain with multiple models of production and a number of distinct players. Figure 1 shows a simplified illustration of where contract manufacturing situate in the industry value chain.

Chip design and manufacturing are downstream from R&D and upstream from distribution of chips to Original Equipment Manufacturers (OEM) and Original Design Manufacturers (ODM).[2] Abstractly, chip design and manufacturing can follow one of two models of production. The first model is the Integrated Device Manufacturer (IDM) model. IDMs design, manufacture, assemble, and test semiconductor chips on their own. They then sell the finished chips to the OEMs and ODMs. Before the 1980s, all semiconductor firms had intergrated design and manufacturing. The "foundry" business originates from these firms buying and selling excess capacity from other firms to even out their capacity needs (McLellan 2012).

[2] R&D in chip design and manufacturing usually involves significant public-private partnerships, for example throught research institutes like SEMATECH. OEMs and ODMs purchase manufactured chips from chip designers. Using those chips, OEMs design and produce electronics based on the their customers' requests while ODMs design and produce electronics either to be used in their own products or to be sold under their own brand names.

Figure 1 **Contract manufacturing in semiconductor industry value chain**



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

A second model of production is the fabless-foundry model originating in the 1980s. A "fab" is short for a semiconductor fabrication plant. Fabless firms design chips and outsource chip manufacturing. As the capital requirements to build manufacturing capacities have increased significantly with newer and newer generations of manufacturing technologies, going fabless has helped many small innovative chip designers to enter the downstream market for chips. However, for fabless firms, securing spare capacity at IDMs was uncertain and risky. In response to the fabless firms' need to address capacity uncertainty, "pure-play" contract manufacturers (or foundries) came into existence with the founding of the Taiwan Semiconductor Manufacturing Company (TSMC) in 1987. When a fabless firm orders chips from a contract manufacturer, the contract specifies the quantity, technical specifications, price, delivery timeframe, and so on. Pure-play foundries do not have their own chip design businesses and exclusively manufacture chips. Outsourced semiconductor assembly and test (OSAT) firms provide third-party packaging and test services for the chips manufactured by the contract manufacturer. The finished chips are then sold to the OEMs and ODMs by the firm who designed the chips.

The distinctions we have described between different links in the value chain are not strict, and a semiconductor firm can occupy more than one position in the value chain. Notably, an IDM faced with capacity constraints can place manufacturing orders to foundries, and an IDM faced with excess capacity can take outside orders and perform contract manufacturing. An IDM, for example Samsung, may even have dedicated capacity for contract manufacturing. One main difference in manufacturing between an IDM and a foundry is that IDM fabs are typically optimized for a limited range of products that the IDM sells downstream, whereas foundries handle a diverse product mix. For example, an Intel fab typically produces only Intel microprocessors, while a TSMC fab could manufacture microprocessors, system-on-chips 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. Among the other leading firms, UMC and SMIC consistently rank in the top eight, accounting for 9 – 21% and 5–8% of the top-eight revenue during this period, respectively. The remaining positions fluctuate due to entry, mergers, and other industry dynamics.[3] TSMC has been by far the largest and the most important player in this market. As such, statistics on TSMC and actions taken by TSMC are often reflective and representative of the overall market conditions in contract semiconductor manufacturing.

2.2. The Product and Manufacturing Technology

A unit of product being manufactured is a silicon wafer with layers of integrated circuits fabricated at its surface. Each wafer carries many chips that will later be cut and packaged appropriately. See Figure 2a for an example of a wafer and Figure 2b for the circuits patterns fabricated on the wafer under a microscope. The most prevalent manufacturing process is the complementary metal-oxide semiconductor (CMOS) process, which uses light and a series of complex physiochemical reactions to transfer intricate circuit designs onto the wafer layer by layer. Our study will only focus on manufacturing capacity for the CMOS process, which accounts for 92% of the orders in our representative demand data.[4]

Figure 2 Semiconductor wafer as the product

(a) Former Intel CEO Pat Gelsinger holding a wafer

(b) Wafer under microscope

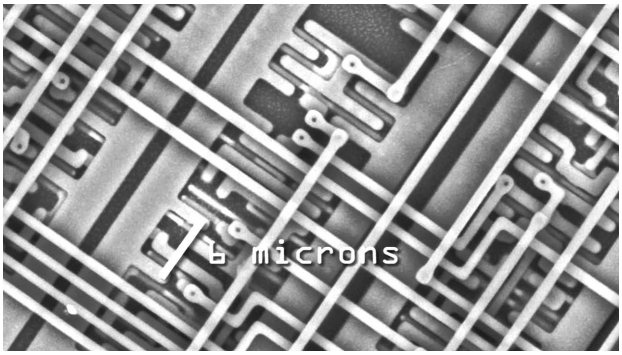


Photo credit: Figure 2a from Shankland (2022) and Figure 2b from NISENet (2014).

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. Until the generation 28nm which

[3] 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.

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

entered commercial production around 2012, these numbers carry exact meaning and measures a specific physical dimension of the chip. A smaller technology 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. After 28nm, though shrinking the integrated circuits becomes increasingly challenging, manufacturers continue 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 a numeric convention, but those 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 has become to think of it 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 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 that technical specifications of chips manufactured using the new node with the same number start to not match from one manufacturer to the next. As of now, node names are more often considered a marketing term.

Semiconductor manufacturing is a prime example of high-tech industries characterized by substantial fixed costs and rapid technology turnover. Building a fab costs billions, and every two years, new manufacturing technology generations gain revenue share at the expense of older ones, as seen in Figure 3a. This equilibrium outcome results from both demand and supply factors. High-value chips like microprocessors require the most advanced technology, and when a leading manufacturer like TSMC introduces a new generation of technology, few competitors initially offer the same. Over time, more rivals climb the quality ladder and enter the market, leading to price competition and driving down profits. As manufacturing technologies constantly advance and become increasingly expensive, manufacturers build fewer, larger fabs for the latest tech to maximize scale economies, as shown in Figure 3b. Consequently, it is crucial for the sustained well-being of manufacturers to gain access to the widest possible market worldwide for each fab of newer generations.

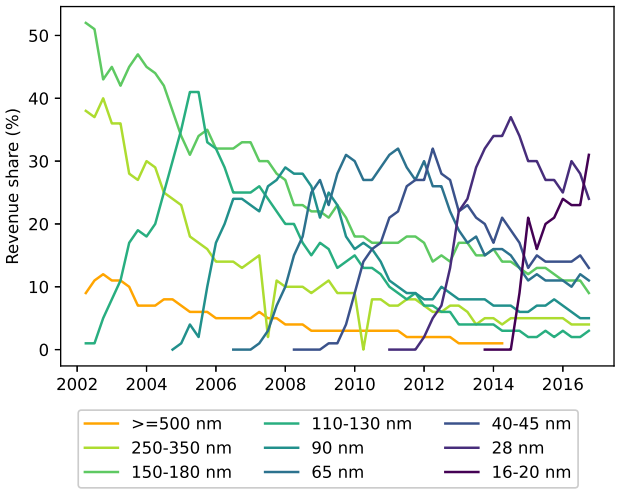
2.3. Policy Background

Recent policies in the U.S. and China on the semiconductor industry distinguish themselves from national policies on the industry in the past by their much larger scale in direct subsidies and by their explicit focus on domestic manufacturing. While government support in the semiconductor industry has a long history and is widespread, the majority of budgetary support for the industry from national governments has historically been focused on R&D activities and tax concessions, which can be found in China, Europe, Israel, South Korea, Taiwan, and the U.S. just to name a few (OECD 2019).

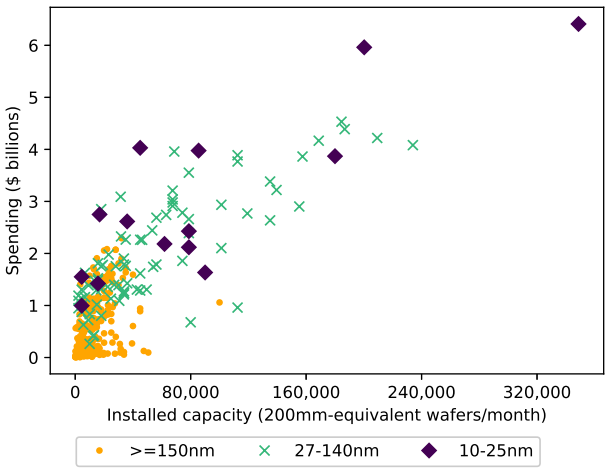
The CHIPS and Science Act of 2022 is a landmark legislation in the U.S. representative of the recent policy development in the U.S. regarding the seminductor industry. The Act provides $52.7 billion to cover an umbrella set of incentives and support around R&D, workforce development and private-public partnerships of the industry, but the centerpiece of the legislation is the CHIPS for America Fund, which includes $39 billion in financial assistance over five years to build, expand, and modernize domestic semiconductor manufacturing (U.S. Senate Committee on Commerce, Science, and

Figure 3 Technology turnover and scale economies of semiconductor manufacturing

(a) Revenue shares of different technologies at TSMC



(b) New fab construction and equipment spending



Notes: Data source for Figure 3a is TSMC quarterly reports. Figure 3b 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.

Transportation 2022a). In addition to the CHIPS Fund, the Act also created a 25% investment tax credit (ITC) for investments in semiconductor manufacturing (U.S. Senate Committee on Commerce, Science, and Transportation 2022b). Moreover, both the CHIPS Fund and the ITC also restrict recipients from building advanced semiconductor production facilities in countries that "present a national security concern," defined as China, Iran, Russia, and North Korea, for ten years after receiving an award (Swanson 2023). The U.S. Senate Committee on Commerce, Science, and Transportation cites the fact that only 12% of chips are currently manufactured domestically, compared to 37% in the 1990s, and many foreign competitors, including China, are investing heavily to dominate the industry as the main motivations for the Act. In addition, the U.S. has imposed semiconductor export controls on China (Khan 2020, Industry and Security Bureau 2020, 2022, 2023). These include list-based restrictions that limit the sale of specific advanced manufacturing equipment and chips to Chinese entities, requiring exporters to secure licenses. Moreover, end-use and end-user controls target specific Chinese entities, notably major chip producers including SMIC. Spanning a broader set of technologies, these controls generally assume a default position of denying export licenses.

Before the passage of the CHIPS and Science Act, China indeed has significantly surpassed the U.S. in providing national-level incentives targeted at semiconductor manufacturing. While the Chinese government has long had policies to support its nascent chip industry (VerWey 2019), these efforts accelerated in 2014, when China released its the Guidelines to Promote the National Integrated Circuit Industry (IC Guidelines), which laid out ambitious targets for its domestic semiconductor industry (Xinhua News Agency 2014, Semiconductor Industry Association 2021). A year later, China published

the Made in China 2025 Plan, which set aspirational goals for China to achieve 70% self-sufficiency in semiconductors by 2025 (The State Council of the People's Republic of China 2015, Semiconductor Industry Association 2021). In contrast to the U.S. which had no federal level grants, subsidies, or tax incentives specifically targeted at semiconductor manufacturing between 2000 and 2020, China has set up in September 2014 the National Integrated Circuits Industry Development Investment Fund (known as the "Big Fund") which was renewed in 2019 for a second round of state financing. The two rounds of financing is totaled at $50 billion (Semiconductor Industry Association 2020). By mid 2021, the Big Fund has invested around $40 billion, of which 70% has been for manufacturing (Semiconductor Industry Association 2021) and the subsidies have been limited to Chinese-owned companies building domestic manufacturing capacities (Zhao 2021, Wang 2022). In addition to the funds, the Chinese national government has announced tax incentives for manufacturers operating at advanced technology nodes. For instance, manufacturers using technology nodes of 28nm and below receive a 10-year corporate tax exemption, while those with nodes of 65nm and below get a 5-year exemption (Ministry of Finance of the People's Republic of China et al. 2020). Reports have also suggested that Chinese manufacturers can obtain loans at sub-market rates from state-owned financial institutions, although the process and funding amount are opaque (OECD 2019).

Other countries and regions, including South Korea, Singapore, Japan, Taiwan, and Europe all have provided a range of grants, subsidies, tax incentives, loans, and other support for semiconductor manufacturing (Semiconductor Industry Association 2020). But the scale of their government support are significantly dwarfed by the recent policy provisions in the U.S. and China. For example, Taiwan has provided below $1 billion in manufacturing grants and subsidies while South Korea has provided an estimate of $7 billion to $10 billion in manufacturing grants and subsidies between 2000 and 2020 (Semiconductor Industry Association 2020).

While the U.S. semiconductor industry largely supported and welcomed the subsidies (Semiconductor Industry Association 2022), the restrictions on investment and sales of chips in China have been a subject of heavy lobbying from the industry, which collectively earns about one-third of its revenue from China. Chip makers 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 also confirm that economics of scale and volume of business from the entire global market are crucial for the health and growth of the industry. The conflicting incentives between national governments and the industry motivates this 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 3–5 years. Recent developments reflect this volatility. On China's side, the third phase of the Big Fund was announced in 2024 (Reuters 2024) with $48 billions of additional financing. In the U.S., the "big

beautiful bill" proposes increasing the investment tax credit (ITC) for semiconductor investment to 35% (CNBC 2025). In addition, ongoing discussions about semiconductor tariffs, which have persisted since early 2025, are likely to continue, with proposed rates reaching 25% or higher (York and Durante 2025). This continuous flux in policy not only creates a highly uncertain environment, but also makes it infeasible to isolate the effects of individual policies using an event-study framework. This motivates our use of a structural estimation and counterfactual approach to analyze the combined impact of policy bundles and firms' strategic responses under different scenarios.

2.4. TSMC Fab Location Case Study

To anchor our examination of how different policy bundles influence major manufacturers' location choices and performance, 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 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 → 90nm → 80nm → 65nm → 55nm → 45nm → 40nm • Eventually settled as mature node facility (40nm b/w 2008-2015) **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 EC.1.

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 has 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.

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 EC.1.8. All monetary values are in USD and are not inflation-adjusted.

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. The rest of notes of Table 2 apply.

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. 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.

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.

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 the 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 an English auction. Each manufacturer's marginal cost decreases as installed capacity increases. Though simple, our model effectively 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:

$$\max_{j \in \mathcal{J}_{igt}} v_{igt} + \kappa_{igt} \cdot \mathbf{1}(geog_{igt} \in \{geog_{jgt}\}) - p_{j,igt}. \tag{1}$$

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 $\mathbf{1}(geog_{igt} \in geog_{jgt})$ 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:

$$c_{j,igt} = \underbrace{c_{gt} + c_j - \zeta \cdot ln(Q^j_{gt})}_{c_{jgt}} + c^{geog}_{j,igt} - \lambda \cdot \mathbf{1}(geog_{igt} \in \{geog_{jgt}\}) + \omega_{j,igt}.$$

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 supply of capacity at manufacturer j for generation g in period t. The ζ term approximates how marginal costs decrease with manufacturer capacity. This term 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_{jgt}.

$c_{j,igt}^{geog}$ 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 regional 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 Given a time period t, before all negotiations begin, buyers choose the technology generation g and the quantity of chips q_{igt}. We treat q_{igt} as exogenous.[7] $I_{gt} \sim D_{gt}(\cdot)$ represents the number 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 i seeks to contract manufacturing capacity through the negotiation process with all manufacturers with available capacity.

We specify a full information negotiation game, that is, 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.

The term $v_{igt} + \kappa_{igt} \cdot \mathbf{1}(geog_{igt} \in \{geog_{jgt}\})$ represents the reservation price. Manufacturers participate in bidding for the buyer's contract only if they can provide positive utility; that is, when $v_{igt} + \kappa_{igt} \cdot \mathbf{1}(geog_{igt} \in \{geog_{jgt}\}) \geq c_{jgt} + c_{j,igt}^{geog} - \lambda \cdot \mathbf{1}(geog_{igt} \in geog_{jgt}) + \omega_{j,igt}$. 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:

$$p_{igt}^{(1)} = v_{igt} + \kappa_{igt} \cdot \mathbf{1}(geog_{igt} \in \{geog_{jgt}\}).$$

[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] It is 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.

Multiple Manufacturers Bidding: If there are multiple manufacturers with available capacity participating, the buyer runs a standard English procurement auction. The winning manufacturer, which we index as $j=(1)$, charges:

$$p_{igt}^{(1)} = c_{(2),gt} + c_{(2),igt}^{geog} - \mathbf{1}(geog_{igt} \in \{geog_{(2),gt}\})(\lambda + \kappa_{igt}) + \omega_{(2),igt} + \kappa_{igt} \cdot \mathbf{1}(geog_{igt} \in \{geog_{(1),gt}\}),$$

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).

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

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

If there are multiple (>1) manufacturers with available capacity bidding, the expected profit of the winner (conditional on winning) per unit quantity ordered is

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

and each manufacturer j wins with probability

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

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

$$E[p_{igt}^{(1)}(\boldsymbol{s}_{gt})] = \sum_{k \in \mathcal{J}_{igt} \setminus j} \left(\frac{exp(-(c_{kgt} + c_{k,igt}^{geog} - \mathbf{1}(geog_{igt} \in \{geog_{kgt}\})(\lambda + \kappa_{igt}))))}{\sum_{j' \in \mathcal{J}_{igt} \setminus j} exp(-(c_{j'gt} + c_{j',igt}^{geog} - \mathbf{1}(geog_{igt} \in \{geog_{j'gt}\})(\lambda + \kappa_{igt}))))} \right.$$
$$\left. \times E[(c_{kgt} + c_{k,igt}^{geog} - \mathbf{1}(geog_{igt} \in \{geog_{kgt}\})(\lambda + \kappa_{igt}) + \omega_{k,igt} + \mathbf{1}(geog_{igt} \in \{geog_{jgt}\})\kappa_{igt}) | k=(2)] \right)$$

The order-level expected profit for j contracting with i is the quantity q_{igt} multiplied by the expected profit upon winning, $E[\pi_{j,igt}(\boldsymbol{s}_{gt})]$, 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}:

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

5. Model Estimation

5.1. Parameterization, Identification, and Calibration

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:

$$c_{gt} = \alpha_0 g^{\alpha_g} t^{\alpha_t}. \tag{2}$$

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]

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 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 $\tilde{c}_{IDM}$.

$c^{geog}_{j,igt}$ takes 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 $c^{geog} = \tilde{c}^{geog} \cdot c_{gt}$. $\tilde{c}^{ROW}$ is normalized to zero. These terms are identified by changes in revenue shares of manufacturers that operate facilities in multiple regions when their capacity composition across regions changes.

ζ captures the decrease in marginal cost as a manufacturer's installed capacity for a technology generation increases. This term is parameterized 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, helps identify ζ.

$\omega_{j,igt}$ captures variation in prices not explained by generation-period or firm-specific means. σ_ω is identified by these remaining price variations.

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

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

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 in different regions. We do not have data on margins for any manufacturer in North America, so we combine North America and ROW into one variable. The overall cost is the sum of facility-specific costs.

In principle, buyer willingness to pay (v_{igt}) could be identified when only one manufacturer participates in bidding, as the equilibrium price would equal the buyer's willingness to pay plus any regional manufacturer benefit (κ_{igt}). However, we do not observe monopoly situations in our data where only

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

[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.

a single contract manufacturer is available for specific technology generations and periods.[10] Additionally, since we lack visibility into which manufacturers actually negotiate with each buyer, we do not impose participation constraints based on reservation prices. Instead, we allow all manufacturers with available capacity for each technology generation and period to participate in bidding, ensuring multiple bidders in all cases. With multiple manufacturers participating, v_{igt} cancels out in the expressions for equilibrium prices, probability of winning, and profit. Therefore, this term remains unidentified, and we do not parameterize it.

We assume that $\kappa_{igt} = c_{gt} \cdot \tilde{\kappa}_{igt}$, where $\tilde{\kappa}_{igt}$ is drawn from an exponential distribution $H(\cdot) \equiv Exp(\sigma_\kappa)$, and $\lambda = c_{gt} \cdot \tilde{\lambda}$, where $\tilde{\lambda}$ is a constant parameter. $\tilde{\lambda}$ is identified by differences in the revenue share from Chinese buyers between mainland Chinese and Taiwanese manufacturers. σ_κ is separately identified by variation in this share across periods and manufacturers.

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.

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

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

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

[10] While TSMC was often the only pure-play foundry capable of manufacturing the most advanced technologies, Intel and Samsung remained technologically competitive with TSMC during our sample period. The competitive landscape has since shifted, with TSMC surpassing its competitors, suggesting that access to more recent data could enable identification of this parameter.

[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 IC designers a quarter before these are reflected in IC 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.

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.[12]

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 a pure-play foundry, all of its capacity is available for contract manufacturing, so its total capacity for contract manufacturing at facility k is Q_{jt}^k. For an IDM, we assume that a random proportion $m_{jt}^k \sim M(\cdot)$ of each facility's capacity is available for contract manufacturing in period t, so the total capacity for contract manufacturing at facility k is $m_{jt}^k Q_{jt}^k$. m_{jt}^k is equal to $min\{max\{0, \tilde{m}_{jt}^k\}, 1\}$, where $\tilde{m}_{jt}^k$ is drawn from a normal distribution $N(0, \sigma_{\tilde{m}})$. IDM facilities' capacity is, in expectation, fully utilized by IDMs themselves for their end products, so the distribution has a mean of zero. We calibrate $\sigma_{\tilde{m}} = 0.05$, as semiconductor fabs usually need to have a utilization above 95% to justify the fixed cost investment. The $\tilde{m}$ term is calibrated, rather than estimated, as it cannot be separately identified from the IDM-specific marginal cost $\tilde{c}_{IDM}$.

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:

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

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 $\boldsymbol{\theta}$.

$\bar{m}$ 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

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

Table 4 Moment Conditions

Moment	Description	Parameters Identified
$margin_{jt}(\boldsymbol{\theta})$	The gross margin of manufacturer j in period t, for the subset of firms with available margin data.	$\{C_{geog}\}$
$pmean_{geog,gt}(\boldsymbol{\theta})$	The average price for each technology generation, period, and fab region.	$\alpha_0, \alpha_g, \alpha_t$
$pvar_{geog,gt}(\boldsymbol{\theta})$	The variance of prices within each technology generation, period, and fab region.	σ_ω
$revenue_{jy}(\boldsymbol{\theta})$	Annual revenue for manufacturer j, available for a group of leading contract manufacturers, including both pure-play foundries and IDMs.	$\{\tilde{c}_j\}, \{\tilde{c}^{geog}\}$
$revenue_{jt}(\boldsymbol{\theta})$	Quarterly revenue for manufacturer j, available for a subset of pure-play foundries.	$\{\tilde{c}_j\}, \{\tilde{c}^{geog}\}$
$revenue_{y}(\boldsymbol{\theta})$	Annual contract manufacturing market revenue, available for a subset of years.	σ_D
$revshare_{IDM,y}(\boldsymbol{\theta})$	The revenue share coming from IDMs among all contract manufacturers, available for a subset of years.	$\tilde{c}_{IDM}$
$utilization_{jt}(\boldsymbol{\theta})$	The average capacity utilization rate for manufacturer j, available for a subset of firm-quarters.	ζ
$revshare_{jt,CN}(\boldsymbol{\theta})$	Revenue share from Chinese buyers for manufacturer j, available for a subset of firm-quarters.	$\sigma_\kappa, \tilde{\lambda}, l_0^{CN}, l_1^{CN}, l_2^{CN}$
$revshare_{jt,NA}(\boldsymbol{\theta})$	Revenue share from North American buyers for manufacturer j, available for a subset of firm-quarters.	$l_0^{NA}, l_1^{NA}, l_2^{NA}$

Note: For information about data availability, please see Appendix EC.1.

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.

6. Estimation Results

Estimation results are presented 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 2, 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 is small 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, such as larger pure-play foundries or IDMs, the estimates can be interpreted as follows: for example, $\tilde{c}_{TSMC} = -0.1396$, $\tilde{c}_{TW} = -0.0940$, and $\tilde{c}_{NA} = 0.0003$. This implies that, on average, a TSMC fab located in Taiwan has a marginal cost that is $[(-0.1396 - 0.0940) - 0.0003] \times 100\% = -23.39\%$ lower than that of a small pure-play foundry in North America, all else being equal. Additionally, a TSMC fab located in Taiwan has a marginal cost that is $[(-0.0940) - 0.0003] \times 100\% = -9.43\%$ lower than if the same fab were located in North America, all else being 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.

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.

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.

6.1. 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 policy counterfactuals concerning manufacturers' payoffs under various geopolitical arrangements, which are the primary focus of this paper. However, they may be important to address for other topics in this setting, such as the overall societal welfare effects of industrial policies, which we plan to explore in future work.

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.[13] 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.

Moreover, our model does not impose a hard capacity constraint. Instead, we assume that the marginal cost of manufacturing declines as installed capacity increases, which induces manufacturers with larger capacity to win more orders and fill their capacity. Under this model, a manufacturer could receive orders far exceeding its installed capacity, whereas in reality, utilization never exceeds 110%. One way to address this is by using a fixed-point technique to solve for the equilibrium price, ensuring that capacity matches the model-implied order quantity at the equilibrium price. Additionally, we have assumed that buyers cannot substitute a different technology generation. This assumption holds largely because substituting a more advanced technology is not cost-effective, and the technical specifications of the chips often prevent the use of more mature technologies. We have also assumed that buyers' demand is inelastic, meaning that the quantity of chips they demand and the number of buyers in the market are unaffected by prices. Future work could relax this assumption by endogenizing quantity with respect to price.

[13] 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.

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. We specify how different policy scenarios alter the parameter values $\boldsymbol{\theta}$ governing the competitive contracting process between manufacturers and buyers relative to our model estimates $\widehat{\boldsymbol{\theta}}$. We then 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 then 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 EC.2 for details).

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.

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). Both are lump-sum payments and do not affect the marginal cost of production or the marginal

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

	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	$0.5B
relocated to the U.S.			[6.7,8.8]	[68.5,76.9]		
Difference (B−A)			−$1.2B	−$1.8B	+$0.8B	+$0.5B
			[−1.8,−0.6]	[−2.5,−1.1]		

Notes: Actual investment refers to the undiscounted sum of investments in Fab 14 from 2000 to 2015. Simulated investment covers the simulated undiscounted sum of investments in Fab 14 for the same period. Actual profit refers to the undiscounted sum of profits for all of TSMC from 2004Q2 to 2015Q3. Simulated profit covers the simulated undiscounted sum of profits for all of TSMC for the same period. We do not isolate Fab 14's profitability due to potential substitution effects across production facilities. The reported simulated profit value is the mean value simulated across 50 sets of $\widehat{\boldsymbol{\theta}}$ estimates, as described in the notes of Table 5. The 90% confidence intervals, shown in square brackets, are based on the 5th and 95th percentiles of these simulations. The reported simulated investment value is the mean predicted value from 30 independently trained MLP neural network models using different random seeds. The corresponding 90% confidence intervals, also in square brackets, are based on the 5th and 95th percentiles of these predictions. For U.S. subsidies, we consider both the direct subsidy and the Investment Tax Credit (ITC), which are the primary manufacturing incentives under the CHIPS and Science Act. The direct subsidy is treated as a lump sum. Since TSMC Arizona is receiving a $6.6 billion direct subsidy on a $65 billion investment, we set this lump sum at 10% of total investment. The ITC currently provides a tax credit equal to 25% of investment, which we also treat as a lump-sum subsidy. In calculating the ITC benefit, we apply a federal corporate income tax rate of 21% and a state rate of 5% (Arizona's rate is 4.9%).

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. A 10% direct subsidy on investment (roughly equivalent to the support received by TSMC's Arizona fabs, $6.6 billion out of a $65 billion total, or 10.2%) and a 25% ITC together amount to $1.3 billion in total incentives for locating the fab in the U.S. While these subsidies roughly offset the higher investment cost, they are insufficient to compensate for the profit losses due to higher marginal production costs. Based on our estimates, relocating Fab 14 to the U.S. in response to capacity subsidies alone—at levels comparable to those offered under the CHIPS and Science Act—would not be a profitable strategy.

C. Fab 14 in Taiwan vs. the U.S. Under U.S. Export Controls In this simulation, we implement export controls on contracts with mainland Chinese buyers, modeled after those in 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 (with 10/12nm and 5/7nm nodes in between). Accordingly, in our export control simulations, we prohibit TSMC from contracting with mainland Chinese buyers for technology generations within four generations of its frontier technology in each period. All other aspects of the competitive environment remain unchanged, and we continue to use cost and demand parameters estimated from the observed data, $\boldsymbol{\theta} = \widehat{\boldsymbol{\theta}}$.

The first column of Table 7 reports TSMC's total profit from 2004Q2 to 2015Q3 under export controls when Fab 14 is located in Taiwan ($67.8B) versus relocated to the U.S. ($66.3B). Overall, export controls reduce TSMC's profitability regardless of Fab 14's location. Compared to the no-policy baseline, TSMC's profit declines by $6.1B[14] when Fab 14 remains in Taiwan. Relocating Fab 14 to the U.S. leads to an additional profit loss of $1.6B.

Notably, the magnitude of the additional profit loss from relocating Fab 14 to the U.S. is similar with or without export controls, suggesting that the effects of relocation and export restrictions are largely additive. Restrictions on selling advanced semiconductors to Chinese customers reduce the pool of buyers TSMC can serve. It turns out that the customers lost due to these export restrictions would have contracted with TSMC regardless of whether it operates from Taiwan or the U.S., resulting in a similar level of profit reduction; thus, there is little additional customer loss attributable to the interaction between export controls and relocation.

Once again, relocating Fab 14 to the U.S. under export controls would not be a profitable strategy, even with capacity subsidies at levels comparable to those provided under the CHIPS and Science Act.

D. Fab 14 in Taiwan vs. the U.S. Under Chinese Tax Exemptions In this simulation, we examine the effects of Chinese tax exemptions. The policy described in Ministry of Finance of the People's Republic of China et al. (2020) reduces the standard 25% corporate income tax rate on profits for chipmakers located in mainland China to a range 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 can be treated 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 (net of region-specific manufacturing costs and benefits).

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: in the Chinese tax exemption simulations, we 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 from the observed data.

[14] $74.0B in Table 6, column 4, minus $67.8B in Table 7, column 1. 90% confidence interval: [5.5, 6.8].

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

	U.S. Export Controls	CN Tax Exemptions	**(Current Regime)** 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: The U.S. export control policy issued in 2022 restricted exports of technology nodes at 16/14 nanometers or below (Industry and Security Bureau 2022), at a time when TSMC's most advanced technology in volume production was 3nm. This reflects a four-generation gap (with 10/12nm and 5/7nm in between). We thus prohibit TSMC from contracting with mainland Chinese buyers for technology generations within four generations of its most advanced technology in each period under export control simulations. Tax exemptions in China for semiconductor manufacturers (Ministry of Finance of the People's Republic of China et al. 2020) reduce the standard 25% corporate income tax rate on profits to a range between 0–12.5%, depending on a firm's technology portfolio and the number of years since the policy announcement. These exemptions can be viewed as marginal cost subsidies for fabs operating in mainland China under our model. Because it is unclear how much of the tax savings are passed on to buyers, we apply a conservative calibration for the CN tax exemption simulations, reducing the region-specific marginal cost in China (c_{CN}) by 5% of the technology-generation- and time-specific common marginal cost (c_{gt}). A tariff rate of 25% or higher has been discussed by the U.S. government (York and Durante 2025). Tariffs can be viewed as a percentage levy on U.S. buyers purchasing imported goods, thereby reducing their willingness to pay when the manufacturer is located outside the U.S. Since this reduction should apply to the total price—not just the profit margin—its effect should be stronger than a 25% corporate tax reduction. However, since the exact effect is uncertain, we conservatively reduce U.S. buyers' willingness to pay (v_{igt}) by 10% of the common marginal cost for fabs located outside of the U.S. under the tariffs simulation. The remaining notes from Table 6 apply.

The second column of Table 7 reports TSMC's total profit from 2004Q2 to 2015Q3 under Chinese tax exemptions, comparing scenarios where Fab 14 is located in Taiwan (\$70.6B) versus in the U.S. (\$68.8B). Overall, the Chinese tax exemptions reduce TSMC's profitability regardless of Fab 14's location. Compared to the no-policy baseline, TSMC's profit declines by \$3.4B[15] when Fab 14 remains in Taiwan. Relocating Fab 14 to the U.S. results in an additional profit loss of \$1.8B.

Interestingly, the magnitude of the additional profit loss from relocation is similar with or without the Chinese tax exemptions, suggesting that the effects of relocation and Chinese tax exemptions are largely additive. The policy-induced reduction in production costs in China weakens TSMC's cost competitiveness, leading it to lose contracts. The types of contracts TSMC loses under the Chinese tax exemption regime appear to be similar regardless of whether Fab 14 is located in Taiwan or the

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

U.S.; thus, there is little additional loss attributable to the interaction between China's marginal cost subsidies and the relocation.

Once again, relocating Fab 14 to the U.S. under marginal cost subsidies in China would not be a profitable strategy, even with capacity subsidies at levels comparable to those provided under the CHIPS and Science Act.

E. Fab 14 in Taiwan vs. the U.S. Under U.S. Export Controls and Chinese Tax Exemptions (Current Regime) In this simulation, we combine the export controls described in Section 7.C with the Chinese subsidies outlined in Section 7.D. To our knowledge, these are the primary policies currently in effect that alter the competitive landscape of the contract semiconductor manufacturing market. We therefore label this simulation the current regime.

The third column of Table 7 reports TSMC's total profit from 2004Q2 to 2015Q3 under this current regime, comparing scenarios in which Fab 14 is located in Taiwan ($65.0B) versus in the U.S. ($63.4B). Overall, the current regime reduces TSMC's profitability regardless of Fab 14's location. Compared to the no-policy baseline, TSMC's profit declines by $9.0B[16] when Fab 14 remains in Taiwan. There is a modest interaction effect—TSMC's total profit loss under the combined policy bundle is slightly smaller than the sum of the losses from the individual effects of export controls ($6.1B) and Chinese tax exemptions ($3.4B). This is because some of the contracts TSMC loses under the combined policy regime would have been lost due to either export controls or Chinese marginal cost subsidies alone. These are primarily high-end Chinese customers affected by export restrictions who could alternatively source from subsidized domestic manufacturers.

Relocating Fab 14 to the U.S. leads to an additional profit loss of $1.6B, which again does not appear to interact meaningfully with either export controls or Chinese tax exemptions. So where does this additional loss come from? Figure 4a provides insight. 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 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

[16] $74.0B in Table 6, column 4, minus $65.0B in Table 7, column 3. 90% confidence interval: [7.8, 10.2].

technologies and a growing share of buyers coming from outside the U.S. 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 TSMC continues its typical fab strategy—building cutting-edge facilities, upgrading them several times, and then letting them mature—it may need to reassess whether this model is sustainable for its U.S. operations.

F. Adding U.S. Import Tariffs to the Current Policy Regime One additional policy option that has recently attracted significant attention is tariffs. A tariff rate of 25% or higher has been discussed by the U.S. government (York and Durante 2025). In this simulation, we examine the potential impact of tariffs in addition to the existing policy regime. 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 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.

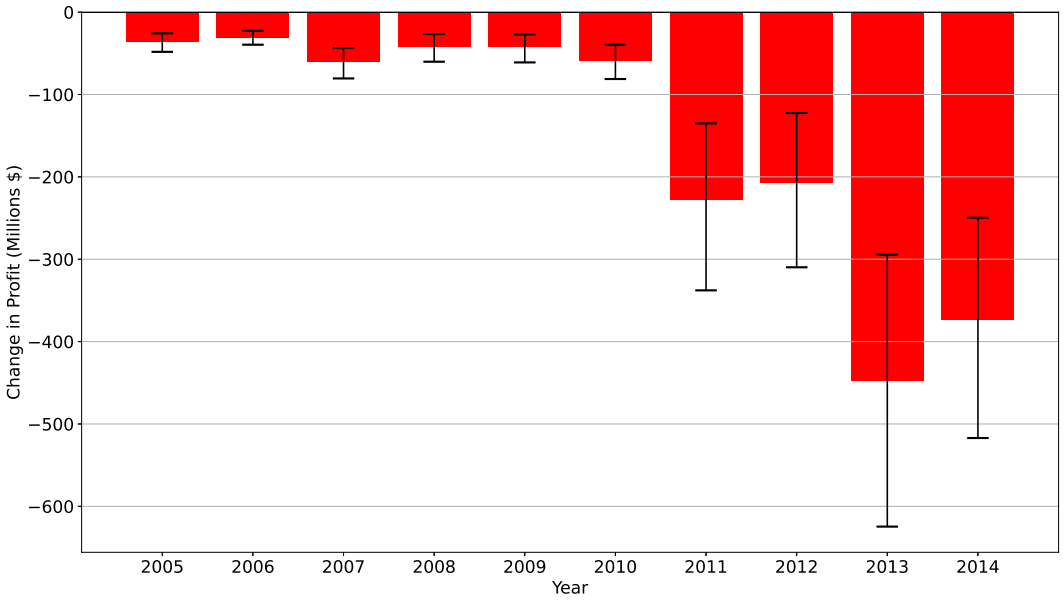
The results are revealing in several respects. First, tariffs have a limited effect when Fab 14 is located in Taiwan. Compared to the current policy regime, adding tariffs results in no visible change in TSMC's profit under our baseline 10% calibration. Even increasing the calibration to 20% or 25% leads 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. 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.[17] 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 4b 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

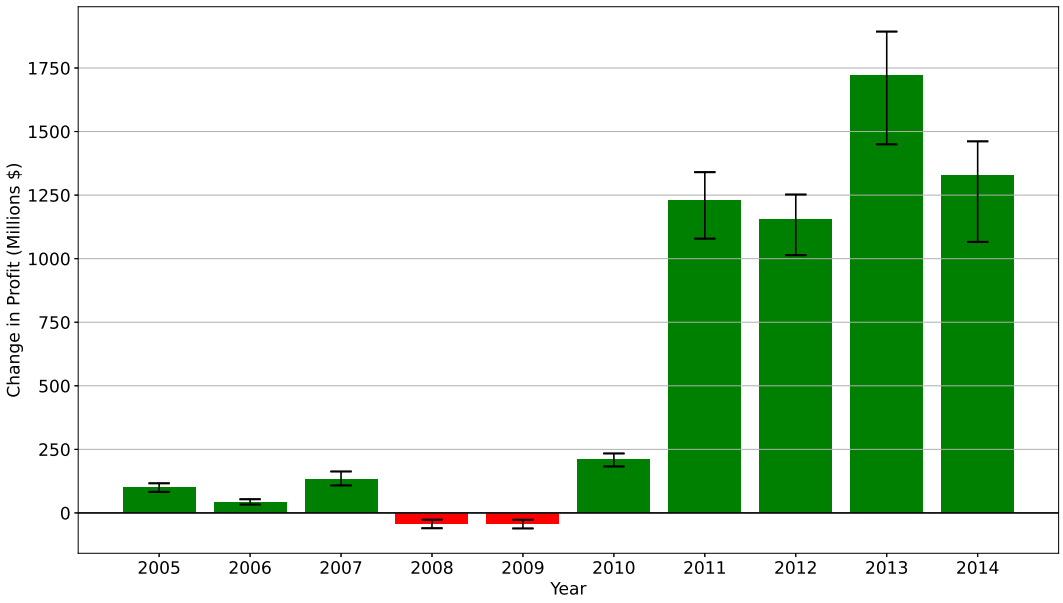
[17] However, downstream demand could be more elastic for high-end products such as flagship smartphones or standalone GPUs used in gaming or High-Performance Computing (HPC) clusters, where the semiconductor share of system value is higher.

Figure 4 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.

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. Tariffs enable the U.S. fab to secure this demand, offsetting profit losses stemming from export controls and Chinese tax exemptions. In

this context, tariffs effectively counterbalance the competitive disadvantages typically faced by fabs producing mature technologies in high-cost locations like the U.S., leading to the opposite patterns observed in Figure 4b compared to Figure 4a.

In summary, our analysis shows that under a tariff scenario, relocating Fab 14 to the U.S. becomes a profitable strategy. Even under the modest tariff rate simulated here, the additional U.S. contracts and $5.9 billion in profit from a domestic presence more than offset the additional $1.2 billion investment cost.

References

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

Allen J, Clark R, Houde JF (2019) Search frictions and market power in negotiated-price markets. *Journal of Political Economy* 127(4):1550–1598.

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

Bollinger B, Gerarden T, Gillingham K, Vollmer D, Xu DY (2024) Strategic avoidance and welfare impacts of solar panel tariffs Working Paper.

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

Chang SJ, Matsumoto Y (2022) Dynamic resource redeployment in global semiconductor firms. *Strategic Management Journal* 43(2):237–265.

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

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

Eizenberg A (2014) Upstream innovation and product variety in the US home PC market. *Review of Economic Studies* 81(3):1003–1045.

Fontana R, Greenstein S (2021) Platform leadership and supply chains: Intel, Centrino, and the restructuring of Wi-Fi supply. *Journal of Economics & Management Strategy* 30(2):259–286.

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

Fuchs E, Kirchain R (2010) Design for location? The impact of manufacturing offshore on technology competitiveness in the optoelectronics industry. *Management Science* 56(12):2323–2349.

Funke M, Wende A (2022) Modeling semiconductor export restrictions and the US-China trade conflict URL `https://www.ssrn.com/abstract=4307050`, working paper.

Gardete PM (2016) Competing under asymmetric information: The case of dynamic random access memory manufacturing. *Management Science* 62(11):3291–3309.

Goettler RL, Gordon BR (2011) Does AMD spur Intel to innovate more? *Journal of Political Economy* 119(6):1141–1200.

Goldberg PK, Juhász R, Lane NJ, Forte GL, Thurk J (2024) Industrial policy in the global semiconductor sector NBER Working Paper 32651.

Gugler K, Siebert R (2007) Market power versus efficiency effects of mergers and research joint ventures: evidence from the semiconductor industry. *The Review of Economics and Statistics* 89(4):645–659.

Hall BH, Ziedonis RH (2001) The patent paradox revisited: an empirical study of patenting in the US semiconductor industry, 1979-1995. *Rand Journal of Economics* 101–128.

Han P, Jiang W, Mei D (2024) Mapping U.S.–China technology decoupling: Policies, innovation, and firm performance. *Management Science* mnsc.2022.02057.

Hayashi Y, Fitch A (2023) U.S. tightens curbs on AI chip exports to China, widening rift with Nvidia and Intel. *Wall Street Journal* URL `https://www.wsj.com/tech/u-s-tightens-curbs-on-ai-chip-exports-to-china-widening-rift-with-u-s-businesses-3b9983df`, accessed on January 1, 2024.

Henderson R (1993) Underinvestment and incompetence as responses to radical innovation: Evidence from the photolithographic alignment equipment industry. *The RAND Journal of Economics* 248–270.

Hruska J (2021) How are process nodes defined? *ExtremeTech* URL `https://www.extremetech.com/computing/296154-how-are-process-nodes-defined`, accessed on January 4, 2024.

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

Igami M (2017) 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.

Igami M (2018) Industry dynamics of offshoring: The case of hard disk drives. *American Economic Journal: Microeconomics* 10(1):67–101.

Igami M, Kusaka S, Qiu J, Tran TL (2024) Welfare gains from product & process innovations: The case of LCD panels, 2001–2011 .

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

Industry and Security Bureau (2020) Addition of entities to the entity list, revision of entry on the entity list, and removal of entities from the entity list. *Federal Register* URL `https://www.federalregister.gov/documents/2020/12/22/2020-28031/addition-of-entities-to-the-entity-list-revision-of-entry-on-the-entity-list-and-removal-of-entities`, accessed on March 18, 2024.

Industry and Security Bureau (2022) Implementation of additional export controls: Certain advanced computing and semiconductor manufacturing items; supercomputer and semiconductor end use; entity list modification. *Federal Register* URL `https://www.federalregister.gov/documents/2022/10/13/2022-21658/implementation-of-additional-export-controls-certain-advanced-computing-and-semiconductor`, accessed on January 1, 2024.

Industry and Security Bureau (2023) Implementation of additional export controls: Certain advanced computing items; supercomputer and semiconductor end use; updates and corrections. *Federal Register* URL `https://www.federalregister.gov/documents/2023/10/25/2023-23055/implementation-of-additional-export-controls-certain-advanced-computing-items-supercomputer-and`, accessed on January 1, 2024.

Irwin DA, Klenow PJ (1994) Learning-by-doing spillovers in the semiconductor industry. *Journal of political Economy* 102(6):1200–1227.

Jacobs BW, Singhal VR, Zhan X (2022) Stock market reaction to global supply chain disruptions from the 2018 US government ban on ZTE. *Journal of Operations Management* 68(8):903–927.

Kapoor R (2013) Persistence of integration in the face of specialization: How firms navigated the winds of disintegration and shaped the architecture of the semiconductor industry. *Organization Science* 24(4):1195–1213.

Khan SM (2020) U.S. semiconductor exports to China: Current policies and trends URL `https://cset.georgetown.edu/publication/u-s-semiconductor-exports-to-china-current-policies-and-trends/`, accessed on March 18, 2024.

Kim D, VerWey J (2019) The potential impacts of the Made in China 2025 Roadmap on the integrated circuit industries in the U.S., EU and Japan URL `https://www.ssrn.com/abstract=3433844`, working paper.

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

McLellan P (2012) A brief history of semiconductors: the foundry transition. URL `https://semiwiki.com/semiconductor-manufacturers/1611-a-brief-history-of-semiconductors-the-foundry-transition/`, accessed on January 3, 2024.

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

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), Ministry of Industry and Information Technology of the People's Republic of China (2020) 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. URL `https://www.gov.cn/zhengce/zhengceku/2020-12/17/content_5570401.htm`, accessed on December 31, 2023.

NISENet (2014) Zoom into a microchip. URL `https://www.youtube.com/watch?v=Fxv3JoS1uY8`, accessed on January 4, 2024.

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

Park DJ, Liu S (2023) A study on the economic effects of U.S. export controls on semiconductors to China. *Journal of International Trade & Commerce* 19(1):129–142.

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

Salomon R, Martin X (2008) Learning, knowledge transfer, and technology implementation performance: A study of time-to-build in the global semiconductor industry. *Management Science* 54(7):1266–1280.

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

Semiconductor Industry Association (2016) Beyond borders: the global semiconductor value chain 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 (2020) U.S. needs greater semiconductor manufacturing incentives. URL https://www.semiconductors.org/wp-content/uploads/2020/07/U.S.-Needs-Greater-Semiconductor-Manufacturing-Incentives-Infographic1.pdf, accessed on December 30, 2023.

Semiconductor Industry Association (2021) SIA whitepaper: Taking stock of China's semiconductor industry 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 (2022) SIA applauds House passage of CHIPS Act, urges President to sign bill into law. URL https://www.semiconductors.org/sia-applauds-house-passage-of-chips-act-urges-president-to-sign-bill-into-law/, accessed on December 30, 2023.

Shankland S (2022) Fear not: Intel says the chip technology that'll power your PC in 2025 is going great. URL https://www.cnet.com/tech/computing/intel-shows-off-the-chip-technology-thatll-power-your-pc-in-2025/, accessed on January 4, 2024.

Swanson A (2023) U.S. issues final rules to keep chip funds out of China. *New York Times* 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 (2020) Tsmc announces intention to build and operate an advanced semiconductor fab in the united states. URL https://pr.tsmc.com/system/files/newspdf/THGOANPGTH/NEWS_FILE_EN.pdf.

Terwiesch C, Ren ZJ, Ho TH, Cohen MA (2005) An empirical analysis of forecast sharing in the semiconductor equipment supply chain. *Management Science* 51(2):208–220.

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

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

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

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

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

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

Varas A, Varadarajan R (2020) How restrictions to trade with China could end US leadership in semiconductors 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.

Varas A, Varadarajan R, Goodrich J, Yinug F (2020) Government incentives and US competitiveness in semiconductor manufacturing 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.

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

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

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

York E, Durante A (2025) Tracking the economic impact of the Trump tariffs. *Tax Foundation* URL https://taxfoundation.org/research/all/federal/trump-tariffs-trade-war/, last updated July 17, 2025; Accessed July 21, 2025.

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

Appendices

EC.1. Data Appendices

EC.1.1. 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.

EC.1.2. 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

Table EC.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 EC.1.4)	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 EC.1.5)	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 EC.1.6)	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

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

EC.1.3. 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).

EC.1.4. 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.

EC.1.5. 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.

EC.1.6. 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.

EC.1.7. 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).

EC.1.8. 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.[18]

[18] We note that in recent years, as shrinking the transistors has become increasingly challenging, node names have lost their connection to the physical features on the chip. However, node names of newer generations still follow this numeric convention, 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.

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:

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

EC.2. Predicting Investment Costs from Characteristics of Capacity Investment Projects

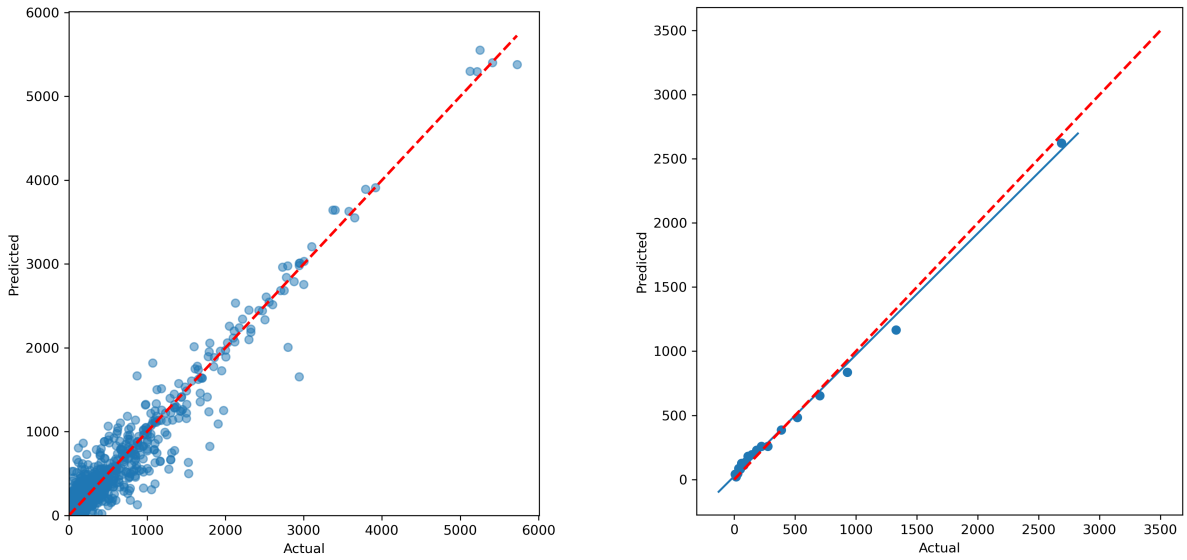
We fit multilayer perceptron neural networks to predict the investment amount of project inv_{jt}^{k}. We use the following project characteristics to predict these outcomes:

- 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.

Figure EC.1 Actual values and in-sample predicted values for investment amount

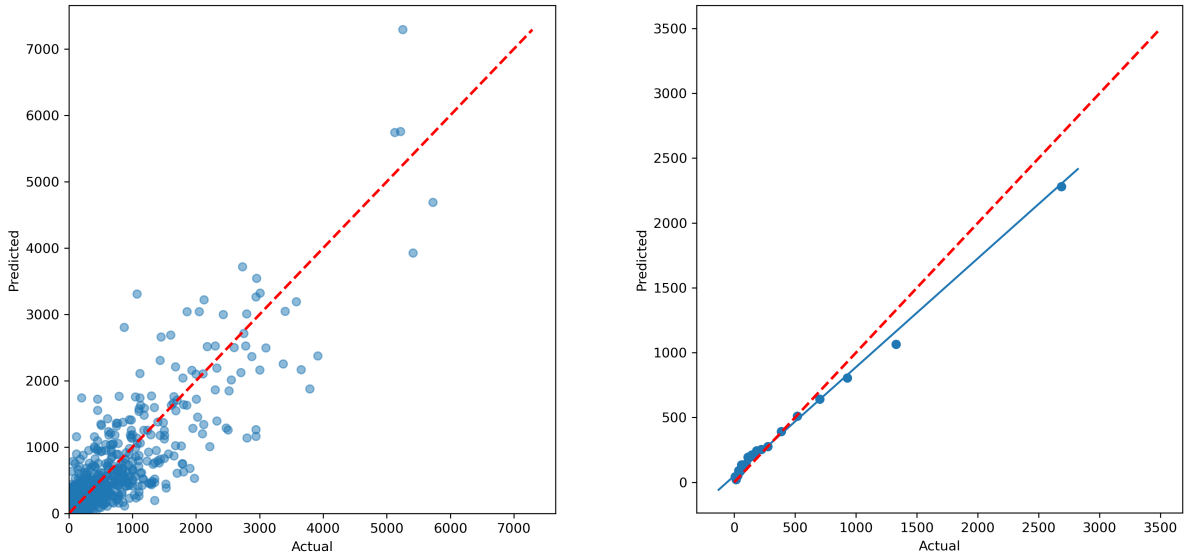
(a) Scatterplot (b) Binned scatterplot



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.

Figure EC.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) (2024a) Chinese Yuan Renminbi to U.S. Dollar spot exchange rate [DEXCHUS]. 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) (2024b) Taiwan Dollars to U.S. Dollar spot exchange rate [DEXTAUS]. URL https://fred.stlouisfed.org/series/DEXTAUS, retrieved from FRED, Federal Reserve Bank of St. Louis, Februrary 5, 2024.

Bower N (2016) Top 13 IC foundries by foundry sales in 2015. *Electronic Specifier* URL https://www.electronicspecifier.com/news/analysis/top-13-ic-foundries-by-foundry-sales-in-2015.

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

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

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

Design & Reuse (2020a) China forecast to represent 22% of the foundry market in 2020 URL https://www.design-reuse.com/news/48819/china-foundry-market-in-2020.html, accessed on February 5, 2024.

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

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

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

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

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

Mutschler AS (2008) Pure-play foundries comprise 84% of market, IC Insights says. *Electronic Design News (EDN)* 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 (2012) Apple pushes Samsung's foundry sales yet again! URL https://siliconsemiconductor.net/article/75826/Apple_Pushes_Samsungs_Foundry_Sales_Yet_Again, accessed on January 28, 2024.

US Bureau of Labor Statistics (2024) 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, URL https://fred.stlouisfed.org/series/CPIAUCSL.

Acknowledgments

We are very grateful to Shane Greenstein, Myrto Kalouptsidi, Robin Lee, Ariel Pakes, and Elie Tamer for their time, patience, and insightful advice. We would also like to thank Ashley Chang, Renee James, Willy Shih, and David Yoffie for thoughtful discussions about the semiconductor industry. We thank David M. Byrne, Ed Hall, and Shungo Saito for data-related inquires. We thank participants of the Harvard Industrial Organization Workshop, the Ross Technology & Operations Brown Bag, and the Columbia Business School IEOR-DRO Seminar for their helpful suggestions. We thank the Doctoral Office at Harvard Business School for financial assistance. All errors are our own.