

No Shortage of Profit:
Technological Change, Chip ‘Shortages’, and Capital Accumulation
in the Semiconductor Business

Christopher Mouré

A Major Research Paper submitted to the Department of Politics in partial fulfillment of the
requirements for the degree of Master of Arts in Political Science

Supervised by: Professor Jonathan Nitzan

Second reader: Professor Jennifer Pybus

York University

Toronto, Ontario

August 2022

Table of Contents

ABSTRACT.....	3
SECTION 1: INTRODUCTION	4
1.1 MAIN ARGUMENT	4
1.2 OUTLINE OF PAPER.....	6
SECTION 2. THEORETICAL FRAMEWORK	7
2.1 NEOCLASSICAL THEORY	8
2.2 MARXIST POLITICAL ECONOMY	14
2.3 THE CAPITAL AS POWER VIEW: SUBJUGATING CREATIVITY TO POWER	17
2.3.1 <i>Differential accumulation</i>	18
2.3.2 <i>Power and strategic sabotage</i>	20
2.3.3 <i>Accumulation strategies: breadth and depth</i>	21
SECTION 3. METHODOLOGY	22
3.1 MEASURING ‘DOMINANT SEMICONDUCTOR CAPITAL’	22
3.2 CONSTRUCTING THE BENCHMARK	25
3.3 MEASURING ‘SHORTAGES’	25
3.4 MEASURING ‘BREADTH’	28
3.5 NOTE ON QUALITATIVE EVIDENCE FROM SECONDARY LITERATURE.....	29
SECTION 4. ORIGIN AND DEVELOPMENT OF THE US SEMICONDUCTOR BUSINESS	30
4.1 MILITARY AND GOVERNMENTAL ORIGINS.....	30
4.2 ANTI-TRUST AND COMPULSORY LICENSING	33
SECTION 5. NO ‘SHORTAGE’ OF PROFIT: AN EMPIRICAL STUDY OF DOMINANT SEMICONDUCTOR CAPITAL	37
5.1 QUANTITATIVE ANALYSIS.....	37
5.1.1 <i>Power, production, and pricing</i>	38
5.1.2 <i>Differential profit and the perception of shortage</i>	42
5.2 THE 1986 US-JAPAN SEMICONDUCTOR TRADE AGREEMENT	45
5.2.1 <i>The Semiconductor Industry Association and the creation of a business-government distributional coalition</i>	47
5.2.2 <i>Pressuring Japan</i>	50
5.2.3 <i>The ‘shortage’ of 1987-1990</i>	51
5.2.4 <i>The privatization of price controls</i>	52
5.2.5 <i>The uncanny accuracy of production ‘forecasts’</i>	53
5.2.6 <i>The entrance of Samsung and the second STA</i>	54
5.2.7 <i>Summing up</i>	55
5.3 A NEW BREADTH MILLENNIUM.....	57
5.3.1 <i>Centrifugal forces</i>	58
5.3.2 <i>Quantitative analysis</i>	61
5.3.3 <i>A return to depth?</i>	67
SECTION 6. CONCLUSION	69
6.1 MUMFORD’S WARNING	70
DATA APPENDIX	71
RESULTS (METHOD 2)	73
BIBLIOGRAPHY.....	76

Table of Figures

FIGURE 1. PERCENT CHANGE IN ‘REAL’ OUTPUT AND PERCENT CHANGE IN PRICES, U.S. SEMICONDUCTOR MANUFACTURING.....	40
FIGURE 2. PERCENT CHANGE IN DIFFERENTIAL RETURN ON EQUITY AND PERCENT CHANGE IN CAPITAL EXPENDITURES, DOMINANT SEMICONDUCTOR CAPITAL (METHOD 1).....	41
FIGURE 3. RETURN ON EQUITY AND SEMICONDUCTOR SHORTAGES, DOMINANT SEMICONDUCTOR CAPITAL AND COMPUSTAT 500 AVERAGE (METHOD 1)	43
FIGURE 4. DIFFERENTIAL PROFITABILITY AND SEMICONDUCTOR ‘SHORTAGES’ (METHOD 1).....	44
FIGURE 5. ANNUAL VOLUME OF ACQUISITIONS BY MONETARY VALUE, SEMICONDUCTOR MANUFACTURING	62
FIGURE 6. SECTOR SIZE AND DOMINANT SEMICONDUCTOR CAPITAL’S SHARE OF THE REVENUE (METHOD 1).....	63
FIGURE 7. SECTOR SIZE AND CAPITALIZATION VOLATILITY, DOMINANT SEMICONDUCTOR CAPITAL (METHOD 1)	64
FIGURE 8. DIFFERENTIAL BUY-TO-BUILD AND DIFFERENTIAL PROFITABILITY, DOMINANT SEMICONDUCTOR CAPITAL (METHOD 1).....	66
FIGURE 9. PERCENT CHANGE IN DIFFERENTIAL RETURN ON EQUITY (PROFIT) AND PERCENT CHANGE IN CAPITAL EXPENDITURES, DOMINANT SEMICONDUCTOR CAPITAL (METHOD 2)	73
FIGURE 10. RETURN ON EQUITY AND SEMICONDUCTOR SHORTAGES, DOMINANT SEMICONDUCTOR CAPITAL AND COMPUSTAT 500 AVERAGE (METHOD 2)	73
FIGURE 11. DIFFERENTIAL PROFITABILITY AND SEMICONDUCTOR ‘SHORTAGES’ (METHOD 2).....	74
FIGURE 12. SECTOR SIZE AND DOMINANT SEMICONDUCTOR CAPITAL’S SHARE OF THE REVENUE (METHOD 2).....	74
FIGURE 13. SECTOR SIZE AND CAPITALIZATION VOLATILITY, DOMINANT SEMICONDUCTOR CAPITAL (METHOD 2) ...	75
FIGURE 14. DIFFERENTIAL BUY-TO-BUILD AND DIFFERENTIAL PROFITABILITY, DOMINANT SEMICONDUCTOR CAPITAL (METHOD 2).....	75

Abstract

Rapid technological change is often touted as a fundamental reality of capitalist societies. It is also often presented as concrete evidence for the supposed progressive improvement of material well-being that characterises the capitalist system of social order. Since its emergence in the mid-20th century, semiconductor technology in many ways exemplifies this reality. Yet the rapid advancement of semiconductor technology has also been accompanied by social conflict. The history of the technology is as much a story of frequent global chip ‘shortages’ and geopolitical disputes as it is one of exponentially growing computational power. The purpose of this study is to examine how the two sides of this story—progress and conflict—are linked. Starting from the theoretical political economic framework of capital as power, I put organized social power at the central of this inquiry. I examine the behaviour of large semiconductor manufacturing firms in an attempt to uncover empirical relationships between capital investment, chip ‘shortages’, prices, and profits. Using quantitative and qualitative analysis, I find that rapid technological change in the business of semiconductors can be both a problem and a solution in the pursuit of differential capital accumulation.¹

¹ According to capital as power theory, differential capital accumulation refers the accumulation of organized social power relative to other capitalist groups. Differential accumulation is discussed further in Section 2.3.1.

Section 1: Introduction

In this project, I examine the relationship between technical change and capitalist power in the US semiconductor business. The theoretical approach used is the capital as power framework, which argues that, in their quest for accumulation, capitalists seek to subjugate human creativity—including the creative and open-ended processes of technological change—to power. Furthermore, the capital as power framework argues that capitalists seek to *differentially* accumulate power – that is, relative to other capitalists—and the primary means of differential accumulation is ‘strategic sabotage’: the measured disruption of social, creative, and cooperative processes. One implication of this argument is that firms developing and selling new technology might accumulate by *limiting* technical change as well as *unleashing* or *propelling* it. In this paper, I engage with these arguments through the study of a business sector famous for the exponential progress of its core technology, integrated circuits.² In studying how semiconductor firms accumulate while at the same time pursuing technical change at breakneck speed, I make a critical appraisal of the thesis that capital accumulation occurs through strategic sabotage. In doing so, I hope to test the relevance of capital as power theory in studying technological change under capitalism.

1.1 Main Argument

The main arguments of the paper can be summarized in five propositions:

Proposition 1. Historically, the rapid pace of technological change in the semiconductor business was partly caused by *limits* placed on the ability of dominant electronics firms to sabotage the creative processes necessary for technical change. Specifically, a combination of military

² This progress is exemplified in the ongoing fulfillment of ‘Moore’s Law’. Moore’s law denotes the general prediction, initially formulated by Gordon Moore in 1965, that the number of transistors that could be fit onto a microchip would double every two years while the minimum costs per transistor would be halved in the same time.

involvement and anti-trust laws preventing the monopolization of semiconductor technology through intellectual property unleashed ‘industry’ and encouraged fragmentation and rapid change over oligopoly and stagnation.³

Proposition 2. As a result of these restraints on business control, firms differentially accumulated primarily by coordinating the strategic limitation of the production volume of new chip technologies as they were introduced. The lasting result of this strategy is that firms have been able to increase their differential profitability through temporarily increased prices, justified by the perception of a chip ‘shortage’.

Proposition 3. Dominant semiconductor firms achieved effective control of semiconductor production capacity primarily through lobbying, business-government coalitions, and implicit cooperation between firms.

Proposition 4. The speed of technological change, while propelled in part by firms’ need to maintain competitiveness, was more of a *problem* for differential accumulation than an aid. This problem was particularly evident from the mid-1990s to the mid-2000s when the growth of new firms increased uncertainty regarding the future earning capacity of dominant firms.

Proposition 5. In the mid-2000s, dominant semiconductor firms responded to the centrifugal forces of technical change through a wave of mergers and acquisitions, resulting in a reduction in the volatility of their capitalization and a significant increase in their differential profitability.

³ The meaning of the term ‘industry’ here is taken from the work of Thorstein Veblen. His concepts of ‘business’ and ‘industry’ are discussed in Section 2.3.

1.2 Outline of paper

The outline of the paper is as follows. In Section 2, I explain the logic behind the choice of capital as power for the starting theoretical framework. I give a brief explanation for why I reject two alternative political-economic theories—neoclassical economics and Marxist political economy—for this study. I then outline the relevant theoretical concepts of capital as power. In Section 3, I describe the research methodology of the paper by introducing some core empirical concepts and procedures. I explain how and why these empirical categories are used and how the notion of ‘shortages’—despite its theoretically problematic definition—might be operationalized into a quantitative, if approximate, measure. In Section 4 I explain the origins and development of the US-based semiconductor business, with a particular emphasis on factors relevant to the sector’s characterization as one of rapid technological change. I make a case for Proposition 1: that the combination of early military and government support during a period of historically strong antitrust enforcement set the stage for a business sector that was structurally limited in its ability to sabotage the pace of technical change.

Section 5 comprises the bulk of the analysis and is divided into three sub-sections. In Sub-section 5.1, I present quantitative evidence to support Proposition 2: the correlation between dominant semiconductor firms’ differential profitability and periodic chip ‘shortages’ indicates an ongoing accumulation strategy of differential inflation (price increases). I show that periods of differential *decumulation*, in which dominant firms trailed average profitability are almost always followed by the appearance of a chip ‘shortage’, and during ‘shortage’ periods dominant firms tend to beat, rather than trail the average. In this context, the perception of shortages is explained as a predictable consequence and convenient justification for this particular strategy of differential accumulation. In Sub-section 5.2, I qualitatively analyze this accumulation process through a

detailed case study of the 1986 Japan-US Semiconductor Trade Agreement (STA), providing evidence for Proposition 3. I explain how, when in the late 1970s, Japanese semiconductor producers began to threaten US control over chip production, US-based firms coordinated in lobbying the US government for help. These firms successfully convinced the US government to pressure the Japanese government and Japanese chip producers to reduce output and collude in raising prices. This section provides qualitative evidence to the broader trends analyzed in Sub-section 5.1 and draws some general conclusions about the qualitative power dynamics of dominant capital in this sector. In Sub-section 5.3, I present quantitative and qualitative evidence for Propositions 4 and 5. I show that in the mid-1990s, the power of dominant semiconductor firms encountered the threat of new firm growth, which had a destabilizing effect. Consequently, firms began an extensive strategy of mergers and acquisitions (M&A) that allowed dominant firms to increase their differential profitability and stabilize perceptions about their future earning capacity.

In Section 6, I offer some concluding thoughts about the broader dynamics of capitalism and technology. I also note that the reader's conclusion should *not* be that social institutions must be redesigned to leave processes of technical change completely unrestrained, or that social conflict can be solved simply by accelerating the pace of technical change. Instead, I suggest that regardless of the pace of change, in present-day capitalism, the invention of new technologies should not be equated uncritically with increased human wellbeing.

Section 2. Theoretical framework

If the purpose of our inquiry is to understand how the logic of capitalism shapes technical change, the theoretical framework chosen should be one that adequately explains that logic in a way that can be measured empirically. Two of capitalism's fundamental organizational concepts

are the price system and the profit motive, and both are difficult if not impossible to understand without quantities. Thus, the ability of a political-economic theory to explain the sources of profit and the movement of prices is a decisive criterion for theory choice. The first part of this section argues that neither neoclassical theory nor Marxist theory adequately explains where profits come from or how prices are determined.

Theories of capital and profit in neoclassical economic theory rely on implausible and unverifiable assumptions about the nature of capital and its ‘contribution’ to the production process. Moreover, the concepts of supply and demand upon which the explanation of shortages is based are unobservable. Marxist political economy is rejected because its unit of value used to explain the source of profit and prices—socially necessary abstract labour time (SNALT)—is also unobservable. Thus, its relation to prices and profits cannot be empirically shown.

The second part of this section outlines the alternative theory—capital as power—that this paper will primarily engage with. Unlike neoclassical or Marxist political economy, capital as power defines capital not as a productive magnitude but as the symbolic quantification of institutionalized social power. In place of supply and demand, or abstract labour, capital as power offers new concepts centred on the notion that capital represents power *over* social production and explains profit and prices as quantified manifestations of that power.

2.1 Neoclassical theory

Central to the neoclassical explanation of prices is the notion of profit as the price or cost of capital (Moseley 2012, 131). In neoclassical theory, capitalist income is conceived as a “cost” of production, and capital is seen as interchangeable with other inputs like labour or raw materials

(Hunt and Lautzenheiser 2011, 290). Income, or the ‘price’ of capital, is argued to be the owner’s compensation for the contribution of his/her capital to the production process.

The question is, what ‘cost’ is incurred to the capitalist for the use of capital? According to neoclassical economics, capitalists suffer the cost of abstinence. As Hunt and Lautzenheiser write, “while it was always clear to all economists that labour was a social cost of production, conservative economists from [Nassau] Senior onward argued that the capitalists’ abstinence involved a similar social cost and a similar disutility” (2011, 289). While some economists, like Alfred Marshall, use other terms such as “waiting,” or “capital services,” the substance of their argument is the same (289). Because capitalists cannot ‘consume’ their capital goods while they are in use in the factory, neoclassical economists argue that income represents ‘just’ compensation for the disutility of abstention. Profit is the capitalist’s reward for abstaining from consumption.

One problem with this theory is that any monetary quantification of suffering incurred by waiting is inherently subjective. Since suffering is a qualitative, relativistic phenomenon, there is no objective method of determining an external measure of suffering. One person’s suffering cannot be experienced by another, and thus it is difficult to discover objectively whether a capitalist has received the correct reward.

In addition, there is also the more serious analytical problem of measuring the material-productive quantity of capital itself. The central problem of measuring capital is that while capital goods are made of heterogeneous unit quantities, capital is valued in a single unit: price. Capital goods comprise physical objects like factories and machinery. Simply adding the physical quantities of these objects together does not give a single coherent ‘value’ of capital. There must first be an intermediate step of transforming physical, heterogeneous entities into a universal material-productive quantity. According to neoclassical theory, the practical solution is to assume

that the quantity of ‘real’ capital is equal to its discounted money price. What this means is that the ‘real’ value of a capital good is taken as the present value of the total earnings that good will generate over the course of its existence. Unfortunately, this assumption introduces a measure of circularity. One needs to know the income that a particular capital good is expected to generate to calculate the discounted money price. However, using the income to derive the quantity of capital assumes what the theory is meant to explain, namely that income is determined by the ‘real’ value of the capital.

Despite the circularity of neoclassical capital theory, the conception of capital as a ‘cost’ of production carries on to this day, albeit with further superficial changes. One update is the assumption that “capital goods are *rented* by producing firms (from the owners of capital), rather than purchased, and thus the supply of capital goods is assumed to be provided by capital goods rental firms” (Moseley 2012, 133-134). This assumption implies that there must be a market for capital goods in which such rental prices are determined. However, because the firm usually owns the capital goods it supposedly rents, there is no such market; thus, in most cases rental prices cannot be shown to exist. In addition, these imagined rental prices are still derived *from other prices* rather than the ‘real’ quantities they purport to express. Thus, they rely on the same unfounded assumption of equivalence between price and ‘real’ value.

In the US, the Bureau of Labour Statistics currently uses this approach. As explained in one of the technical papers on the statistical practices of the BLS, “‘capital services’ [i.e., the prices of capital] are obtained by renting a specified quantity of capital goods for a specified (short run) period at a competitive rental price” (Harper 1997, 7). In theory, the ‘competitive rental price’ is supposed to represent the portion of total marginal productivity attributable to that particular capital good. Yet this representation is empirically irrefutable, because the aggregation problem

makes the underlying ‘real’ productivity of the said capital unobservable. In practice, the economists at the BLS simply assume that profit (rebranded as ‘capital services’ or ‘cost of capital’) is equivalent to the marginal productivity of the capital goods ‘consumed’ (read depreciated) in the production process. The BLS reports:

We usually cannot obtain direct measures of the prices or quantities of capital service flows. However, one important type of data related directly to this "shadow" market is readily available: property income...capital's cost share is the ratio of property income to the nominal value of output. (Harper 1997, 7)

In other words, neoclassical economists simply assume what they need to explain – the source of profits. The *assumption* is that the income/wage accruing to each individual is equal to the marginal productivity of their factor of production. However, this assumption remains empirically unfounded because capital productivity (as an aggregate quantity) cannot be measured in ‘real’ units.

An additional problem relates to the study of shortages. The neoclassical explanation of shortages is straightforward: shortages occur because there is too much ‘demand’ chasing too little ‘supply’. For instance, the chip shortage that arose during the covid-19 pandemic has been described as the result of “a supply-demand imbalance” in that “demand continues to exceed supply” (Flannery 2022). As the co-chief of Samsung Electronics puts it, “there’s a serious imbalance in supply and demand of chips” (White and Hille 2021). Investment U, the educational arm of financial research group The Oxford Club, explains that “the crux of the semiconductor problem” is that “while demand for chips rose...supply remains stagnant” (Kelly 2022). The logic is simple enough. Shortages, by definition, imply that there is not enough of *something*. However, if supply and demand are unobservable phenomena, then there is no way of determining whether or not a shortage has occurred based on these concepts.

According to neoclassical economics, supply and demand are defined as the aggregate desires or intentions of buyers and sellers to buy or sell a certain quantity of goods at a specific price (Bichler and Nitzan 2021, 28-29). All else remaining the same, if the price of a good *increases*, the amount of that good that a consumer desires to buy *decreases*, and thus this desire can theoretically be plotted on a ‘demand curve’. While buyers want the lowest price they can get, sellers want to sell at the highest price. Thus, while supply theoretically grows with price, demand shrinks. Using the finite resources at their disposal, the efforts of the two parties of buyers and sellers to achieve their goals tend to intersect at an ‘equilibrium’ price point, where both parties are satisfied that they have achieved the ‘best’ price (Bichler and Nitzan 2021, 30).

Unfortunately, the concepts of supply and demand (and by implication, the concept of shortage) are of limited use because they are not empirically observable. Two examples will suffice to explain this fact. First, according to the theory, supply and demand can only be empirically known under conditions of perfect equilibrium. Spontaneous equilibrium only occurs under conditions of perfect competition, meaning the capacity of every buyer and seller to affect the price must be assumed to be equally negligible. Yet in reality, perfect competition no longer exists – if it ever did (Nitzan and Bichler 2009, 72-72). Most business sectors are dominated by a few large firms, meaning that oligopoly, not perfect competition, is the norm (73). Since the assumption of perfect competition is patently false, there is no way to know whether prices reflect supply, demand, and/or some other factors.

Second, supply and demand curves typically represent the equilibrium point for *one* commodity, at a single, hypothetical point in time. Within neoclassical theory, this type of analysis is called “comparative statics” (Nitzan and Bichler 2021, 29). Yet, in reality, prices change all the time, and when they do, the changes themselves have numerous effects – effects which reflexively

impact prices in a dynamic and complex feedback loop (29). Not only does a single price change the supply and demand curves of all other commodities, but it may also change the social conditions within which these transactions are posited. For instance, a single change in prices will redistribute income between buyers and sellers. Because income is one of those external factors which must remain the same in order to generate the comparative static model of supply and demand, each price change “will cause the curves themselves to shift, thus violating the assumption of comparative statics” (35).

In the words of Joan Robinson and John Eatwell, a metaphysical statement contains no factual information, which means that it cannot be changed by facts (Robinson and Eatwell 1973, 9). In this sense, if supply and demand are unobservable, then any argument about their effects is purely metaphysical. Moreover, their fundamental unobservability means that they can be used to ‘explain’ any outcome without the possibility of falsifiability – an essential measure of scientific theories.⁴ In practice, this means that when semiconductor prices rise, the claim that this rise is caused by demand exceeding supply (or supply falling short of demand) cannot be disproven. When prices rise, the rise is explained by an “uptick” in demand, and when they fall, it is due to demand “weakness” (Flamm 1996, 220-221). However, these claims are irrefutable without actual supply and demand curves against which to measure such ‘imbalances’ (Bichler and Nitzan 2021, 34). Like the assumption that profit measures the ‘cost’ of capital, the inability to empirically verify such explanations undermines the analytical power of the theory.

⁴ See Popper [1963] 2002, Ch. 1 for a discussion of the importance of falsifiability to scientific inquiry.

2.2 Marxist political economy

Karl Marx developed his political-economic system at a time of sweeping technological and social changes and understood, perhaps more than anyone in his time, the “integrated interdependence” of the fusion of human labour, technical processes, and social domination that characterized the capitalist system (Nitzan and Bichler 2009, 85). Following in the footsteps of classical economists like David Ricardo, Marx sought to place his analysis on a scientific and empirical footing. Like Ricardo, he located the origin of capital accumulation in the labour process (84). He argued that all profit came from surplus value, the value created by *socially necessary abstract labour*, measured in units of *time* (Hunt and Lautzenheiser 2011, 209). According to Marx, the maximum surplus value available for the capitalist to extract is determined by the amount of abstract *labour time* it takes to make the commodity, while the actual surplus value can be determined by subtracting from this maximum the socially necessary abstract labour time (SNALT) of the *labour power* involved in its production (209). While the argument is plausible—most objects of value are products of some form of human labour or activity—and while it may be simple enough to calculate how long it takes a *single* worker to make a *single* commodity, it is not necessarily clear how much *abstract labour* equates to a given monetary value, or why. This section focuses on one particular difficulty: the problem of turning heterogeneous *labour* into homogenous *value*. The insolubility of this problem suggests that Marxist explanations of profit and prices remain without an objective measure of value and are thus foundationally deficient.

The aggregation problem exists for Marxist political economy on both sides of the Marxist “production function.” On the output side, it is unclear how to add many different products into a single quantity. For instance, there are thousands of products within the semiconductor business, each with different characteristics—DRAMs, EPROMs, ASICs, logic boards—and each product

line often contains several generations of chips, each with various capacities and uses but often in simultaneous production in the same factory.⁵ When each of these products has a different qualitative use, it is unclear how that use can be aggregated into one objective quantity of output.

The same problem applies to labour inputs. There are many types of labour, each contributing to the qualitative process of production. Some semiconductor workers, like designers, do very different tasks at different timescales than a foundry technician or marketing consultant. Other individuals may be more or less skilled at an even more granular level, with some positions needing several years of expensive training before employment. Workers may also be diligent or lazy, and thus their productivity may vary significantly from others. It follows from these facts that different workers likely contribute different amounts of ‘value’ to the commodities they participate in creating. In general, Marx argued that highly skilled or complex labour produced proportionately more value than unskilled labour, but subsequent theorists considered his treatment of this issue incomplete.⁶ There is also extensive literature around the empirical calculation of skilled labor, but in empirical work, Marxist economists tend to use the wage bill as a proxy for SNALT (see for instance Morishima and Seton 1961, 210; Wolff 1975, 937; Foley 1982, 42; Shaikh 1994, 80-84; or Cockshott and Cottrell 2005, 312). The problem with this practice is that such a substitution treats prices (in this case, the wage rate) and labour values as equivalent. In doing so, Marxist economists also assume what they are meant to explain: they assume that labour values are proportionate to labour costs (denoted in prices) and then argue that the correlation between labour costs and prices confirms the labour theory of value.

⁵ DRAMs refer to dynamic random-access memory chips, a highly standardized component in most electronic systems. EPROMs are another memory chip, and ASICs are chips designed for more specialized functions.

⁶ See Caligaris and Starosta 2019 for a summary of Marx’s position and subsequent theoretical debates around the question of calculating skilled-labour values. For some theoretical treatments of the issue, see Rowthorn 1974; Duménil 1983; and Caligaris and Starosta 2019.

Perhaps even more troubling to the problem of actually measuring labour values is the issue of joint production. As Nitzan and Bichler note, in “‘joint’ production processes, where multiple inputs jointly produce multiple outputs...such jointness makes it difficult to identify which inputs are ‘responsible’ for which output” (2009, 101). In a process like semiconductor manufacturing, thousands of materials, machine parts, marketing efforts and plenty of automation are laboured upon by a work force spanning several countries. Not only must each process and each input be conceptually separated from all others to distinguish which labour inputs produced which quantities of output, but the very process of quantifying a qualitative process also contains irreducibly subjective accounting decisions. In addition, immaterial elements of the price valuation like intellectual property rights are not produced by labour (because they are legal constructs), and thus any prices which reflect immaterial assets like intellectual property rights cannot be explained by labour values alone. Because any determination of specific labour values depends on being able to separate the input of labour from other inputs, their empirical measurement necessitates subjective assumptions that must be reasonably justified. As it stands, there are no such disaggregated labour input data available.

Without an empirically observable explanation of *what* determines prices and *how*, there can be no empirical inquiry into the relationship between shortages, prices, and profits. For neoclassical economists, the association is self-evident: profits represent the capital’s marginal productivity, and shortages result from a mismatch of supply and demand. Unfortunately, neither of these arguments is founded on observable data. Similarly, Marxist political economy also lacks an adequate theory of price determination because its own unit of value is equally challenging to measure. As Paul Cockshott puts it, though data denominated in prices is abundant:

Unfortunately...we have no such independent data on [labour] values. This would require a fully disaggregated input–output table along with data on labour inputs in hours. When we compute industry values

in practice, our primary data are nothing other than industry costs: we infer values from these costs, with unavoidable error. (2005, 312)

To avoid error, then, this paper sets aside neoclassical and Marxist theory and engages with an alternative theory that places *power* at the centre of its analysis of prices and profits.

2.3 The capital as power view: subjugating creativity to power

Capital as power diverges from neoclassical and Marxian political economy by arguing that profit is not a magnitude of utility or labour time but a manifestation of organized social power. As such, capital is not a *productive* entity but a symbolic representation of the power struggle among different capitalist groups as well as between these groups and the rest of the population, where the ultimate goal of these capitalist groups is the differential accumulation of social power (Nitzan & Bichler 2009, 218).

Building on the work of economist and social critic Thorstein Veblen, Nitzan and Bichler divide society into two distinct yet interdependent spheres: ‘industry’ and ‘business’. According to Veblen, ‘industry’ consisted of the collectively produced knowledge and creative activity of society. ‘Business’, on the other hand, was concerned only with profit. Crucially, ‘business’ used private property law to restrain industry — an act of strategic exclusion that Veblen called ‘sabotage’ (Veblen 1908, 534–536). Following Veblen, Nitzan and Bichler argue that profit results not from the production of social goods but from a firm’s ability to “strategically limit social creativity and well-being” (2009, 261).

Although business “does not and cannot make industry productive,” it “can and does still ‘propel’ it” (Nitzan and Bichler 2009, 226). What does this mean? For Nitzan and Bichler, the logic of capitalism induces “human beings, organizations and institutions into a state of hyperactivity, constantly shaping and restructuring their interactions” (226). Insofar as that

hyperactivity enlarges the scope of human wellbeing and “the inter-subjectively defined ‘good life’” however, “it simply becomes a part of ‘industry’” (226). On the other hand, such ‘propulsion’ is only profitable to the degree that it “interfere[s] with and partly hamper[s]” the creative and cooperative processes of ‘industry’” (226). The result is that, while business can unleash industry to a greater or less extent when it does so, it is always in service of the future-oriented goal of maintaining or augmenting power over the creative processes of industry.

This hyperactive dynamic is central to understanding how capital accumulation can operate through sabotage and, simultaneously, through the rapid propulsion of technological change. For most of the history of capitalism, both rapid technological change and the strategic sabotage of technological change have existed together. It is possible that rapid technological change *in general* is actually an *effect* of the differential struggle for power over technology. In any case, capital as power theory suggests that contrary to the claims of semiconductor firms, the race for ever faster and more powerful microchips is not and cannot be the central goal of those firms, but only a means to the ends of business. The question is, if semiconductor firms constantly attempt to subjugate technological change to the interests of business, does rapid technological change occur because of this fact or *in spite of it*?

2.3.1 Differential accumulation

Capitalists, then, seek to accumulate power. There are many ways to define power, and usually, some definitions appear more useful or accurate than others depending on the context in which the term is used. In the context of capital as power, some researchers have argued that power refers to the ability to influence others to do things they would not otherwise do. As Ulf Martin puts it, “a person’s *social power*...is their ability to overcome the combined, but not

necessarily coordinated, resistance of all other persons with respect to their goals” (2019, 3). Others, like Nitzan and Bichler, have suggested that social power is “confidence in obedience...the certainty of the rulers in the submissiveness of the ruled” (2009, 17). In this paper, power is used similarly. In addition, it is used to denote both latent or ‘potential’ influence and the active or ‘kinetic’ manifestations of that influence. In short, social power is both ‘held’ and ‘exercised’, and the two forms constantly transform into one another, much like potential and kinetic energy. The ‘rulers’ in this context are represented by ‘dominant capital’, “the leading corporations and key government organs” at the centre of capital accumulation (17). ‘Accumulation’ expresses the changing ability of this group to “control, shape, and transform society against opposition” (17).

Because social power is always held and exercised in opposition to other groups, Nitzan and Bichler argue that it must be measured differentially, as a *relative* quantity, instead of in terms of maximization (2009, 18). What concerns capitalists, they propose, is the size of their power relative to other capitalist groups and relative to the underlying population. A common way to do this is to measure manifestations of power, particularly profit and stock price, relative to an average benchmark. The proliferation of financial indexes attests to the importance and commonness of this differential approach. Notably, differential accumulation can occur under conditions of stagnation as well as expansion. If a given firm shrinks 5 percent while the average firm shrinks by 15 percent, that firm has differentially *increased* its social power by 10 percent relative to the average firm.

2.3.2 Power and strategic sabotage

The primary means of accumulation, Nitzan and Bichler propose, is to sabotage society to just the *right* degree – enough to earn differential profits, but not so much that earnings decline (2009, 236-237). This sabotage tends to generate resistance, and as a result, its strategic implementation often takes subtle, obscure, and incremental forms. Generally, the expectation of at least an ‘average’ return on investment is considered ‘normal’ in capitalist societies, and the associated social sabotage in its broadest sense becomes largely invisible (242).⁷ However, because sabotage is a socially negative phenomenon, it must often be *justified* in terms of some external, unavoidable necessity to minimize resistance.

Nitzan and Bichler argue that, in general, strategic sabotage takes two forms. First, business attempts to redirect industry toward more profitable ends. This may entail, for instance, investing in individually owned, fossil-fuel-dependent car transportation over electrified mass transit; or using proprietary pharmaceutical solutions to solve problems that have social or environmental causes that might otherwise be prevented (Nitzan and Bichler 2009, 234). These forms of sabotage are challenging to quantify and often turn on the assumption of counterfactual arguments (arguments about what might have happened if what actually happened had not). Thus, while they remain theoretically and practically important, they often make for a more challenging or indirect empirical inquiry.

The second form of sabotage Nitzan and Bichler outline is the systematic under-utilization of capacity. This limitation is easier to quantify and often yields surprising and counter-intuitive empirical results. This paper will primarily focus on this second form of limitation by examining whether semiconductor ‘shortages’ are possibly a result of this form of strategic sabotage.

⁷ The normalization of profit as a ‘natural’ fact is often expressed in common sayings like “there’s no such thing as a free lunch” and “that’s just the cost of doing business.”

2.3.3 Accumulation strategies: breadth and depth

Two other concepts are essential to our inquiry. Nitzan and Bichler argue that strategies for differential accumulation can be conceptually decomposed into two distinct categories: ‘breadth’ and ‘depth’ (2009, 328). A breadth strategy consists of expanding the size of the organization (the number of employees) faster than the average, which can be achieved internally through ‘green-field’ investment or externally through mergers and acquisitions.⁸ Mergers and acquisitions are the more successful and preferred approaches because they expand a firm’s *share* of production without increasing the overall level production in the sector as a whole (331). In contrast, green-field investment risks undermining differential profitability, mainly if other firms competitively increase their output in response, causing tacit/open collusion to give way to price wars (335).

In contrast to breadth, which focuses on the relative size of the organization, the strategy of depth consists of raising a firm’s relative profit by increasing profit per employee faster than the average (or lowering it more slowly). While accumulation through depth is effective in the short term, Nitzan and Bichler argue that it is riskier in the long term. The problem is that because depth entails greater conflict, it is likely to meet stronger resistance from society and other capitalist groups (2009, 332). In addition, prolonged depth creates the opportunity for competitors to gain market share by setting lower profit margins. As a result, depth strategies are more often temporary fixes, and can result in unintended structural shifts in the landscape of power. This is especially true for the semiconductor business. ‘Shortages’ tend to occur periodically and for short

⁸ ‘Green-field’ investment refers to the internal expansion of production. For instance, this can include investing in new plant, property, and equipment, hiring employees, or expanding research and development efforts.

periods of one or two years. Even so, this systematic depth strategy has brought new competitors into the fold and galvanized structural changes in the business over the long run (see Section 5.2.6).

Section 3. Methodology

In examining quantitative evidence regarding the semiconductor business, I draw on several concepts and procedures developed in the capital as power literature and add a couple of my own. My main source of empirical evidence is financial data relating to profit and investment by semiconductor firms. My sources are the Compustat Capital IQ financial database, accessed through Wharton Research Data Services (WRDS), and the financial database *Mergent Online*. Data are processed using the statistical tools of Microsoft Excel.

3.1 Measuring ‘Dominant Semiconductor Capital’

As noted above, the dominant capital concept is central to studying capital as power. While all capitalists are compelled to differentially accumulate to survive, our inquiry concerns those that succeed and play a dominant role in shaping society. While dominant capital can be understood as an objective social category, measuring the bounds of dominant capital is always partly a matter of interpretation and conjecture. On the one hand, one only has to think of blue-chip indexes like the Nasdaq or the S&P 500 to evoke the idea of dominant capital and the many ways one might define it empirically. Treated as a quantitative abstraction, it appears as an index of the most powerful groups in society, comprising a more or less integrated coalition of increasingly global firms, owners, governments, and other large organizations. On the other hand, any quantitative representation will always inevitably fall short of describing the complexity and dynamism of the underlying reality.

Within the capital as power literature, dominant capital is usually measured as the numerical or statistical grouping of the top firms by revenue or market value (capitalization). It can be defined as either a fixed numerical group (for example, the top ten, fifteen, or one hundred) or a percentage (the top one, five, or ten percent, and so on.) of firms within a given social space. I follow this practice, noting again that such a measure is only approximate and does not necessarily represent a coherent social unit by other criteria. Moreover, because the social space is dynamic and overdetermined by many social relations, the quantitative analysis is accompanied by an attempt to qualitatively characterize the same power relations in their historical and processual specificity. Firms comprising dominant capital may struggle among themselves as much as they protect specific collective interests; the capitalization of a given firm may be underpinned by racial violence, national rivalries, and gender inequality; and resistance to capitalist power—often subterranean and unpredictable—is always present, though difficult to quantify. Nevertheless, if capital, measured differentially, represents capitalists' quantified estimates of those forces, and to a large extent, these estimations guide capitalist behaviour, then there are likely to emerge quantitative patterns that align with the qualitative reality.

In this paper, the concept 'Dominant Semiconductor Capital' is used as an approximation for the most powerful US-listed semiconductor manufacturing firms. To construct this empirical measure, I started with the entire North America Compustat Industrial database from 1970-2020. This database comprises the largest public firms listed on North American stock exchanges. I then isolated all firms with the GIC industry code 45301020, which includes only semiconductor manufacturing firms. Firms that produce semiconductors, but either do so solely for in-house use (for instance, IBM) or as unlisted subsidiaries of a larger corporation (for instance, Samsung) are excluded from this list. In the first case, these producers do not share the same business interests as

merchant producers. They do not profit from the production of semiconductors directly, but only from the sale of goods containing semiconductors. As such, I assume that they do not necessarily share an interest in raising the price of semiconductors through strategic sabotage. In the second case, the issue is more empirical than theoretical. It is often difficult, if not impossible, to determine from the available financial data what portion of a firm's total revenue or income is derived from semiconductor manufacture. As a result, I decided that the risk of contaminating the analysis with unrelated financial data was greater than the benefit of trying to include these firms.

In addition, some of the firms listed are not US-based companies. However, my goal is not to construct a US-specific database but an approximate index of global semiconductor firms that can be easily measured. As I will argue below, from the 1980s onwards, Dominant Semiconductor Capital increasingly organized itself globally, rather than nationally, even if the power of national governments remained a necessary part of the differential power of those firms, and most of the leading members of this groups are listed in the US. A significant omission in this context, due to the scope of the project and the availability of data from these regions, is several Japanese and Chinese firms which would make the dataset more representative of the global state of the business.⁹ While this is an unfortunate limitation, such additions must wait for a later project.

After constructing the 45301020 (semiconductor firm) database, I created two Dominant Semiconductor Capital datasets: one comprising the top ten firms by market capitalization and the other the top 10% of firms by market capitalization.¹⁰ The results between the two do not differ greatly, and so in the interest of readability, I will present my analysis using the first method in the

⁹ In the case of Japan, and for the greater part of the period analyzed, most semiconductor manufacturing firms were also subsidiaries of larger consumer electronics conglomerates.

¹⁰ The list of firms was recalculated for each year based on that year's leading firms.

body of the paper, while results based on the second method are shown in a data appendix at the end of the paper.

3.2 Constructing the benchmark

To create a benchmark against which to measure the differential accumulation of Dominant Semiconductor Capital, I constructed a set of the largest 500 firms by capitalization within the entire Compustat North America Industrial database (reselected annually). I then calculated the weighted average of each relevant data point – revenue, profit, capital expenditures, acquisitions, etc. This benchmark dataset is meant to represent the average large firm, against which dominant semiconductor firms might theoretically judge their own accumulation (similar to a measure like the S&P 500). To calculate the weighted averages, I divided the total revenue (or profit, capital expenditures, etc.) by the number of firms, and used these averages to calculate derived measures, like average return on equity. The benefit of this approach (as opposed to using unweighted averages) is that in a weighted average, the size of individual firms determines the size of their effect on the calculation of the average. Because I am focusing on the behavior of large firms under the assumption that larger firms are large *because* of their distinguishing behavior, giving equal weight to smaller and larger firms could potentially be misleading as to the content and effects of that behavior.

3.3 Measuring ‘shortages’

The theoretical starting point for capital as power is that price changes are manifestations of social power and the ability to use it strategically. In analyzing the role of ‘shortages’ in the semiconductor business, the empirical object is two-fold. First, we must determine if there is a

relation between ‘shortages’, price changes and profits; Second, we must evaluate what qualitative power relations underpin these quantitative relationships.

How do we determine what constitutes a shortage? Unfortunately, this question may, in part, have to remain unanswered. As noted in Section 2.1, the term ‘shortage’ is problematic because it is conventionally defined in terms of neoclassical economic theory. According to neoclassical theory, a shortage is a situation in which, at a given price, the quantity demanded exceeds the quantity supplied. The problem with this formulation is that neoclassical demand and supply denote the desires of buyers and sellers, and these desires cannot be empirically observed. Since no one can tell the difference between these unobservable magnitudes, there is no way to tell the extent of the resulting ‘shortage’. Even if a shortage is described in ‘practical’ terms—for instance, when a greater number of orders is received than can be physically manufactured within a given time frame in a given factory—problems of measurement and definition still arise. For instance, how does one know whether chip users intend to buy the chips they order? During a ‘shortage’, chip users can place multiple orders at different firms and then later cancel some of them, or buy more chips than necessary, and then sell the surplus (Flamm 1996, 233). How does one determine what is real and what is fake ‘demand’ from these possibilities?

Similarly, how does one determine a chip factory’s ‘true’ technical capacity? Factories rarely work at ‘full’ capacity and available capacity utilization measurements do not reflect *technical* capacity utilization but *profitable* capacity utilization (Nitzan and Bichler 2009, 234). A firm’s need to achieve a specific rate of return shapes its perceptions about the correct production level, apart from considerations of the needs of chip users (which, of course, may be just as difficult to measure). Neoclassical economics assumes that full technical utilization and profitable full utilization are equivalent. However, even concrete attempts to define ‘shortage’ still face

theoretical and methodological difficulties if this assumption is not granted. Thus, another methodology is needed to examine the role of ‘shortages’ in the semiconductor business.

One empirically observable phenomenon is the *perception* that there is a shortage. The perception of a shortage is relevant because if the actual existence of a shortage is difficult, if not impossible, to verify, then it is *only* the perception that in the end, *justifies* the raising of prices. Justification is necessary because price increases tend to redistribute income from buyers to sellers (provided the level of quantities exchanged does not drop so low that the redistribution is reversed), and thus rising prices tend to meet resistance. Resistance can be countered by having a good *reason* for raising prices, preferably one that places the blame for higher prices outside the seller’s control. The presumed existence of a shortage fulfills this social requirement, whether or not there is a ‘real’ shortage. It is not necessarily a question of the seller ‘lying’ about a shortage – it can simply be a matter of reducing production or restraining investment in new production capacity until a shortage becomes ‘inevitable’. By measuring the perception of shortages, we can examine its correlation with prices while avoiding the theoretical and empirical problems of measuring the shortage itself, and without losing sight of the relationship between shortages and prices as one of justification in the face of resistance.

In short, the methodological problem of measuring ‘shortages’ is provisionally solved by looking only at the perception of shortages. The paper does this by using public mentions of semiconductor shortages through news publications as a proxy for the existence of a general perception. Using the *Nexis Uni* online database, I made a targeted search to identify all newspaper articles between 1980 and 2020 reporting on a semiconductor shortage in either the headline or the body of the text.¹¹ A list of the headlines and links to the articles were downloaded into

¹¹ The Boolean expression used in the search was < (“semiconductor” OR “chip”) AND “shortage”>, and the search was limited to English language publications.

Microsoft Excel and filtered to find all *headlines* mentioning chip shortages, for a total of 83 articles spanning April 1984 to December 2020. For expediency, results were not counted for 2021, because the first few pages of results returned over one hundred headlines mentioning chip shortages. The annual number of mentions is not relevant to the study – one or more mentions in a year is counted as the existence of the perception of a chip shortage.

It is plausible that greater mentions during a given year indicate either a more widespread perception of shortage, a perception of greater seriousness of shortage, or both. However, this study does not attempt to quantitatively define either widespread-ness or seriousness, labelling years only as having a perception of a shortage or not. Finally, the focus on the perception of shortages over ‘real’ shortages does not exclude the possibility of a technical shortage. If it is the perception that a shortage is occurring that ultimately justifies the price increase, whether or not a shortage is actually occurring is irrelevant to our inquiry.

3.4 Measuring ‘breadth’

In Section 5.3, I turn to a discussion of breadth strategies of accumulation. To analyze the breadth strategies of Dominant Semiconductor Capital, I constructed ‘buy-to-build’ indicators for Dominant Semiconductor Capital and the Compustat 500. This indicator is intended to approximate the relative emphasis on external (M&A) over internal (green-field) breadth. It is calculated as the ratio of acquisition expenditures to new capital expenditures (creation of new plant, property, and equipment). This indicator, first conceived by Nitzan and Bichler, is a novel measurement related directly to capital as power research.¹² As noted in Section 2.3.3, whereas green-field investment tends to become unruly and undermine differential profit, mergers and

¹² For more on the buy-to-build indicator, see Nitzan and Bichler 2009, Ch. 15; and Francis 2013.

acquisitions increase firms' relative size without increasing the total run of production (Nitzan and Bichler 2009, 335). Thus, the logic behind the measurement is that, over time, firms pursuing a breadth strategy will devote an increasing share of investment to acquisitions, and dominant firms, *in particular*, will tend to exhibit a higher ratio of acquisitions to green-field investment than the average firm over time. In Section 5.3.2.2, I further develop this concept by measuring the differential buy-to-build ratio of Dominant Semiconductor Capital compared to the Compustat 500 average.

3.5 Note on qualitative evidence from secondary literature

All sources carry a risk of ideological and personal bias, which applies as much to semiconductor business accounts like any other. Journalists and writers working in a particular business sector can develop close and often friendly relationships with their subjects, writing for years on the same companies and mixing socially in the same circles. In addition, their careers depend on maintaining an 'even-handed' approach that can lead them to soften uncomfortable truths about the business (an approach that may also be the policy or bias of the publishing institution). As a result, journalistic accounts often oscillate somewhere between hyperbole and hagiography. Ideologically, journalistic accounts of business also tend to hew closely to the terms of neoclassical economic theory. This means, for instance, that profits are seen to be a simple manifestation of the creative and industrious contributions of a firm (or even simply the CEO) and that the larger dynamics of a sector tend to be explained with uncritical reference to the forces of 'supply' and 'demand'. Nonetheless, these accounts can be valuable for the intimate, insider details of events, and the authors' direct interactions with business executives and other figures. Thus, such 'insider accounts' are used here cautiously and critically.

Section 4. Origin and development of the US semiconductor business

The following section briefly details the role of the US military and government and the part of compulsory licensing of semiconductor patents in fostering a business sector oriented towards rapid technological advances – creating a relatively open and fragmented business landscape.

4.1 Military and governmental origins

The military began its involvement in the semiconductor business by creating it. During World War II, the military sponsored an extensive research program involving thirty to forty research labs to improve the silicon diodes used in radar (Flamm 1996, 29-30). After the transistor was invented in 1947 in Bell Telephone Laboratories (a participant in the government program), it became clear that the small size, greater reliability, and lower cost of transistors would lead them to replace the vacuum tube amplifiers that were widely used in electronics at the time (30). Military involvement in semiconductor research was directed through the Signal Corps Engineering Laboratory of the US Army (31). It started in 1950 and grew to “20 percent of total [semiconductor research] funding by 1952, and 50 percent of transistor work by 1953...stay[ing] at that level through 1955” (31). Flamm estimates that “about 25 percent of Bell Labs’ semiconductor research budget over the period 1949-58 was funded by defence contracts, and all of the early production of Western Electric, the Bell System’s manufacturing affiliate, went to military shipments” (31). The Signal Corps expanded its research funding beyond Bell Labs in 1955 and doubled the amount of its semiconductor funding to \$1 million a year in 1956 (31). It also gave \$50 million to engineering development between 1952-64, which focused on bringing technology beyond the

prototype stage and into mass production (31). In 1956, “\$15 million in contracts was appropriated, with funds flowing to virtually every semiconductor company in the United States” (31). The Signal Corps also funded the development of photolithography, which would become a critical process in the mass manufacture of integrated circuits (30). In short, the US military funded the research and paid for the resulting products at premium prices (24). Government influence was so widespread that a congressional report estimated that “if university and federal laboratory work were factored in, along with engineering development funds, and indirect R&D funding embedded in contracts to procure new devices at premium prices... the federal government paid for 85% of all US electronics R&D in 1959” (32-33). Because the military was interested in creating a large supply of high-performance transistors, many individual firms were paid to build factory capacity far in excess of orders (33).

The reliance on military and government funding only increased with the invention of the integrated circuit (IC) in the late 1950s.¹³ While the companies that developed ICs rejected military funding for research and development to maintain control over the resulting intellectual property, almost all ICs in the early 1960s were used in defence systems and thus, the military, as virtually the only customer, remained the primary source of funding for IC development (34).¹⁴ Tilton notes that defence production as a percentage of total production was 100 percent in 1962, 94 percent in 1963, and 85 percent in 1964 (Tilton 1971, 91). By 1968, the share was still 37 percent (91). In total, some estimates put federal funding at between 40 and 50 percent of all industrial semiconductor R&D from the late 1950s to the early 1970s (Flamm 1996, 36). While military funding for R&D dropped sharply off in the 1970s, the military remained a critical driving force in developing certain specialized areas of the technology (36). In addition, though the government

¹³ Integrated circuits contain several components, including transistors, on a single discrete chip.

¹⁴ The first major application of ICs was in the Minuteman II guided missile (Flamm 1996, 34).

share of semiconductor consumption declined, one source still estimated that “in the mid-1980s, purchases by defence agencies alone—direct and indirect—accounted for over one-quarter of US semiconductor shipments” (37).

Interestingly, the US military during the 1950s and 60s tended to favour new, untested firms (Flamm 1996, 32). For instance, “new firms (those with no background in the older vacuum tube business) accounted for 69 percent of military sales, and 63 percent of all semiconductor sales” (32). While Flamm argues this predilection was due to the military’s quest for high-performance products, the logic may also have been to undermine differential power and keep the creative forces of ‘industry’ relatively unconstrained (32).

The pervasive influence and role of the US Military had three critical effects on the sector relevant to our inquiry here. First, the military’s eagerness and deep pockets pushed firms to focus on high-performance rather than cost-effectiveness. This meant that the accumulation strategies of semiconductor firms were geared towards advancing the technology as quickly as possible in the hopes of landing a lucrative military contract. Second, military largesse in R&D funding lowered the risk of investment by subsidizing the costs of inventing the technology in the first place. Third, the preference for small firms may have undermined the differential power of established electronics firms, in a bid to ‘propel’ the industrial processes of technological change rather than restrain them. Almost none of the dominant vacuum tube firms became dominant in the US semiconductor business. At the same time, there is evidence that in Europe, the reliance on established electronics firms irreversibly slowed the introduction of semiconductor technology (Flamm 1996, 24-25).¹⁵ In effect, the strategies for accumulation within the semiconductor

¹⁵ See Malerba 1985 for an account of the historical development of the European semiconductor business.

business were heavily shaped by the US military's interests, making rapid technological change an unavoidable priority.

4.2 Anti-trust and compulsory licensing

The second important factor in early semiconductor development in the US was a relatively open system of intellectual property (IP) controls. In David Noble's *America by Design* (1979), he argues that since the enactment of the Patent Act of 1836, the US patent system has vastly augmented the power of large corporations by granting temporary monopolies on new science-based knowledge (84-87). By the 1930s, he writes, it was clear that patents benefitted firms over the inventors themselves, whose control over their activities was arrested by "the compulsory signing away of patent rights of employees" (90). In addition, patent pooling between large firms effectively locked out new entrants (93). The result was a system that many saw as an impediment to the growth of technical and scientific knowledge and the diffusion of valuable technologies.

However, the 1940s brought a dramatic shift in the US Department of Justice's (DOJ) position on antitrust and "a new period of aggressive prosecution of corporate patent monopolies began" (Noble 1979, 88). This general context of strong antitrust enforcement is indispensable for understanding the emergence and development of the semiconductor business. Perhaps the most critical event symbolizing this era was the anti-trust suit brought against AT&T in 1949, culminating in a consent decree in 1956.¹⁶ AT&T's liberal IP policies during and following the suit were instrumental in shaping the structure of semiconductor development (Grindley and Teece 1997, 12).

¹⁶ Both Bell laboratories and Western Electric were subsidiaries of AT&T. Bell Laboratories was the site of research and development, and Western Electric handled manufacturing and licensing.

John Tilton offers a detailed account of the situation in *The International Diffusion of Technology: The Case of Semiconductors* (1971). “During the early fifties,” he writes, “the patent and licensing policies of AT&T were the only ones of importance for a firm aspiring to enter the semiconductor industry...[because] AT&T was the pioneer, held the strategic patents, and possessed the vital know-how” (73). Royalties before the antitrust suit ranged from 0 to 5 percent (74). While Tilton argues that AT&T had not used its patent position to suppress competition, in the 1940s, the Department of Justice was pursuing a relatively aggressive antitrust program, specifically targeting firms holding large patent portfolios (Hart 2001, 928). The antitrust division was led by Thurman Arnold, who took over in 1938 and whose “stated objective was to convert antitrust... into a tool for ‘breaking bottlenecks’, including those that inhibited technological innovation” (928). After his appointment, Hart states that:

Arnold’s dramatic expansion of the use of consent decrees... allowed DOJ to establish the terms for settlement with defendants and excluded the judiciary from the process of resolving many cases. Over the course of the next decade, despite opposition in Congress and from big business and the military, Arnold and his followers moved antitrust policy in an increasingly deconcentrationist direction. Compulsory patent licensing, for instance, for the first time became a common element in antitrust settlements in the immediate post-World War II period. (928)

Under Arnold’s supervision, “[the] DOJ soon filed suit against some of the nation’s best-known high-technology companies, including Standard Oil of New Jersey, DuPont, General Electric, and Alcoa, and focused particularly on the patent holdings of some of these firms” (928).

The AT&T suit was opened in 1949 and ended in 1956 in a consent decree ordering Western Electric, which managed licensing for AT&T, to “license all existing patents royalty-free to any interested domestic firm... and all future patents at reasonable rates” (Tilton 1971, 76). In addition, Western Electric’s semiconductor manufacturing operations were restricted to sales to

the government and AT&T sister companies, severely restricting the future growth of its semiconductor business (68). Even while the suit was ongoing, Tilton argues that it “must have influenced the company’s policy of swiftly disseminating its new technology” (76). For instance, between the opening of the case and the consent decree, Western Electric lowered its maximum royalty rate from 5 to 2 percent and held several symposia in which its representatives exhibited and described the different technologies they had developed, including the transistor, diffusion, oxide masking, and other important developments (74-75).

Another crucial aspect of the consent decree was the stipulation that Western Electric could ask for a cross-licensing provision with its licensees (76). Because Bell Laboratories patents were so central to semiconductor technology, firms had a strong incentive to share their own patents with AT&T to gain access, and the proliferation of cross-licensing agreements only further sped up the process of information dissemination. In addition, due to the tough stance of the DOJ, most other semiconductor firms adopted a similarly liberal licensing approach out of fear of landing their own antitrust suit. In short, combined with the consent decree, the prevailing antitrust atmosphere had the effect of ‘unleashing’ the creative and cooperative forces of ‘industry’ by limiting the sabotage wrought by intellectual property monopolies. It contributed to a much freer flow of information and know-how, making the semiconductor business a more highly fragmented sector than it otherwise might have been.

A final and related factor is the high mobility of engineers, who often leave to start their own firms (Tilton 1972, 78). Bell Laboratories is particularly notable for defections, and “as with patents and licences, Bell Laboratories set a precedent for the industry’s behaviour” (81). Because it could not compete in the commercial semiconductor market and could not offer the same lucrative contracts as other semiconductor firms, it took a more collaborative approach,

maintaining “friendly informal relations” with former employees – compounding the integrative power of its cross-licensing agreements.

To recap: there were two significant reasons for the more collaborative and un-hindered development of semiconductor technology and its rapid advancement and dissemination in the early years of its development. First, the military’s strong influence and funding shaped the efforts of the semiconductor business towards rapid improvements in quality and a focus on high-performance over cost-cutting. Even as military spending as a portion of the semiconductor market shrank in the 1970s, it remained an essential support, both through procurement of high-powered computers and in targeted R&D. Second, the consent decree imposed on AT&T, as well as the broader atmosphere of strong antitrust enforcement against patent monopolies, unleashed industry by making the technological knowledge needed to create and design semiconductors a public resource (Grindley and Teece 1997, 13). The limitations on AT&T also led it to foster a cooperative, integrative approach to semiconductor research. Moreover, the more general atmosphere created by an aggressive antitrust division of the DOJ under Thurman Arnold meant that the consent decree set a precedent that other semiconductor firms would follow. Together, though not exclusively, these factors shaped the structure of the US-based semiconductor business as one focused on high-performance and rapid technological change, often regardless of the cost, and one relatively open to newcomers and to sharing the leading advances amongst one another. Furthermore, this analysis suggests that the ‘unleashing’ of industry was primarily shaped by factors outside and even openly antagonistic to the interests of the business. By restraining the sabotage of industry by business through the carrot (military contracts) and the stick (antitrust enforcement), different organs of the US government and military laid the foundation for a more fragmented business sector that *could not help* but propel technological change rapidly forward in

its attempt to accumulate power. As the next section will attempt to show, this structure proved to be an enduring *problem* for differential accumulation.

Section 5. No ‘shortage’ of profit: an empirical study of Dominant Semiconductor Capital

5.1 Quantitative analysis

This first subsection provides quantitative analysis showing the close relationship between chip production, capital expenditures, and differential profit. It argues that the price of semiconductors is closely and negatively related to the amount of productive capacity made available by semiconductor firms. This means that chip prices tend to increase when production capacity decreases – or, more specifically, chip prices fall more slowly when production grows more slowly. Second, it argues that Dominant Semiconductor Capital has been able to differentially accumulate by restraining—within limits—the production of semiconductors and using the resulting atmosphere of ‘shortage’ to justify raising prices. Inspired by a similar study by Nitzan and Bichler on the role of Middle East conflicts on oil profits, this section shows how ‘shortages’ tend to follow a period of differential *deaccumulation* when Dominant Semiconductor Capital trails the average rate of profitability.¹⁷ These ‘shortages’ also tend to *be followed* by a “reversal of fortune,” in which firms exceed the average (Nitzan and Bichler 2002, 236). In short, firms’ differential earnings are largest when they successfully restrict production to the point of a

¹⁷ In Nitzan and Bichler’s (2002) work, they show how Middle East conflicts have played a similar role to semiconductor ‘shortages’, by causing the perception that oil production is being or will be disrupted and, in turn, justifying the raising of oil prices. Unlike semiconductor ‘shortages’, however, the results of this strategy have been immeasurably more tragic.

perception of a ‘shortage’, which is used as a justification for an ‘depth’ accumulation strategy of differential inflation.

5.1.1 Power, production, and pricing

How can it be empirically shown that this thesis is correct, and that Dominant Semiconductor Capital can differentially profit by limiting chip production? One starting point for gathering data is the US Bureau of Labor Statistics (BLS), which publishes detailed annual production data. Figure 1 uses BLS data to compare the rate of change in US semiconductor production volume and the rate of change of the sector’s price deflator. The price deflator inflation is a proxy for overall price change. It estimates how much of the change in the dollar value of production is a result of ‘pure’ price changes, as opposed to a change in ‘real’ output.

As noted above, due to the problem of aggregation, standard measures of ‘real’ output do not correspond to an actual quantity of goods but to the monetary value of those goods, adjusted to remove so-called ‘pure’ price changes. In addition, there are certain theoretical and methodological problems with determining when a price change is ‘pure’ (i.e., a result of ‘supply’ and ‘demand’ factors) and when it is the result of a difference in the quality of the product.¹⁸ For instance, integrated circuits have relatively short product life cycles, with new chips (that can have different physical characteristics and be intended for different uses) coming on and off the market in as little as 18 months. As Nitzan notes, “whenever the nature of the commodity changes, the measurement of such changes in ‘quality’ is crucial for price and quantity calculations” (1992, 156). However, at the BLS, the methods of dealing with the price/quality question are largely opaque and are approached exclusively from a neoclassical perspective. This presents a problem

¹⁸ See Nitzan 1992, Ch. 5 for a discussion of the price/quality problem.

for researchers critical of the fundamental assumptions of neoclassical theory because “the predisposition of price and quantity data toward the neoclassical economic outlook means that these data may not be altogether suitable to test the neoclassical outlook against competing frameworks” (158). For example, if ‘pure’ price changes are imputed to ‘supply’ and ‘demand’ factors, and this explanation is considered to lack factual information, what does this mean for the validity of the measure of ‘pure’ price change? Nonetheless, even after heavily discounting the meaningfulness of the data, it is interesting that they still appear to confirm the thesis that production and price changes are inversely correlated.

The very tight negative correlation between change in output volume and change in price shown in Figure 1 (-.91, note the inverted right scale) suggests that as the rate of *increase* in production slows, so does the rate of *downward* shift in price. This correlation implies that semiconductor price changes have an inverse relation to changes in production volume.

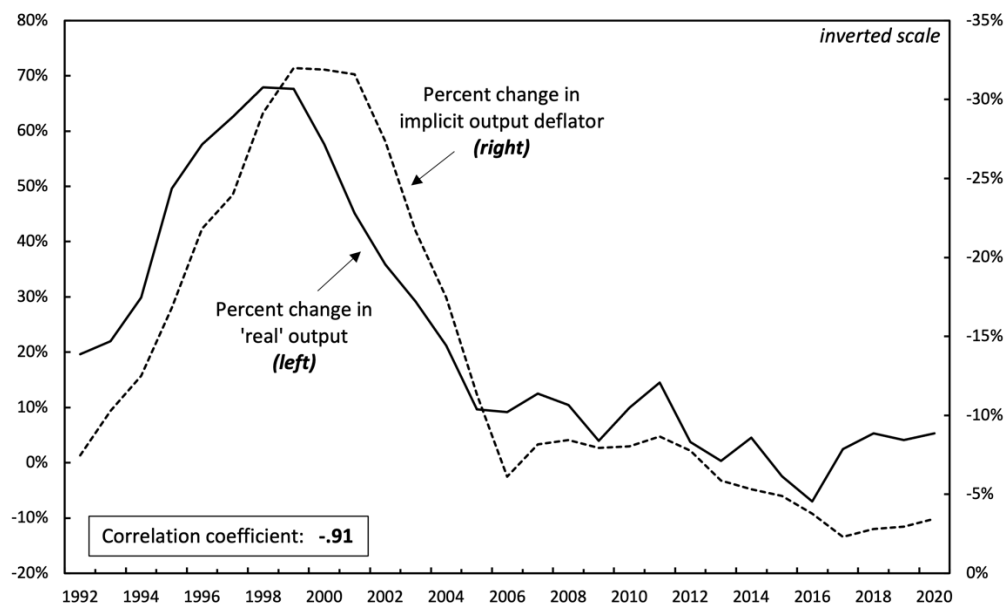


Figure 1. *Percent change in 'real' output and percent change in prices, U.S. semiconductor manufacturing*

If the growth rates of prices and production have an inverse relationship, what about the relationship between differential profit and investment in new production? Figure 2 compares the rate of change of differential profitability to the rate of change in capital expenditures for Dominant Semiconductor Capital (again, note the inverted right scale). Differential profitability is measured here as the percent deviation of Dominant Semiconductor Capital's return on equity (ROE) from the average ROE of the Compustat 500.¹⁹ (From here on, and unless indicated otherwise, the measures for both Dominant Semiconductors and the Compustat 500 are computed as overall weighted group averages.)

¹⁹ The latter measure is calculated as the average return on equity of the largest 500 firms by market capitalization in the Compustat Capital IQ North America database. Return on equity is calculated as net income divided by total common equity.

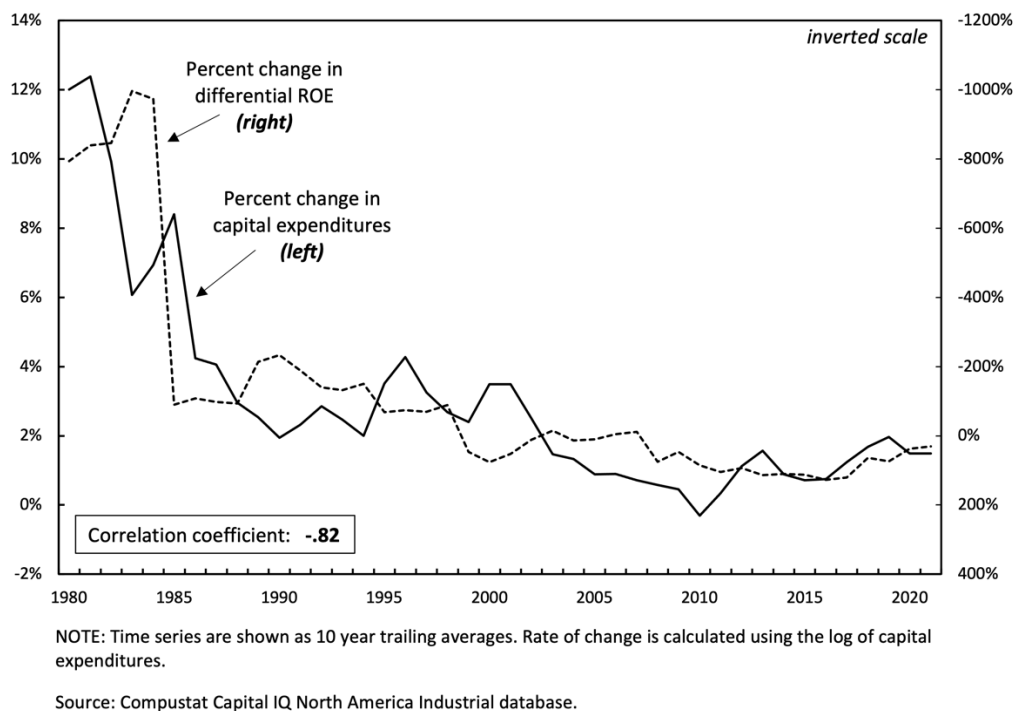


Figure 2. *Percent change in differential return on equity and percent change in capital expenditures, Dominant Semiconductor Capital (Method 1)*

Figure 2 shows a significant negative correlation between the rate of change in differential ROE and the rate of change in capital expenditures (-0.82). The fact that it is negative is counter-intuitive from a neoclassical perspective. If profit is a ‘cost’ of production, logically it should increase faster with a more rapid increase in production, and vice versa when production decelerates. In addition, both profit and production should increase when there is an increase in ‘demand’. Thus, according to neoclassical theory, in a perfectly competitive market the growth of profit and production are likely to move together. However, whereas neoclassical economics focuses on absolute profit growth, in a landscape of shifting prices and antagonistic business relations, what matters is not absolute but *relative* return on investment. What Figure 2 shows is that there is a strong correlation between differential returns and *restraint* on new investment. From a capital as power perspective, this evidence is not as surprising. Capital as power theory

argues that capitalists seek not absolute but differential returns, and that, consequently, they do not primarily seek to expand production but rather to subjugate output to their own differential goals. By strategically limiting production (i.e., by creating more ‘scarcity’), capitalists can charge higher prices for their products. Conversely, if production expands too quickly, they are liable to lose control of pricing, resulting in lower relative prices and lower relative profits, and often lower differential returns. The need to strategically limit production explains the negative correlations shown in Figure 1 and Figure 2.

5.1.2 Differential profit and the perception of shortage

The next stage of the empirical investigation is to look at the relationship between production and profitability in the context of ‘shortages’. If strategically reducing investments in new productive capacity can result in higher prices and higher differential rates of profits for semiconductor firms, it is plausible that ‘shortages’ are not a product of a supply-demand imbalance, as neoclassical theory would have it, but a predictable result of Dominant Semiconductor Capital’s struggle to differentially accumulate.

This next sub-section and Figures 3 and 4 are inspired by Nitzan and Bichler’s work studying the accumulation of arms and oil producers in relation to Middle East wars (see Nitzan and Bichler 2002, Ch. 5). Figure 3 compares the ROE for Dominant Semiconductor Capital and the ROE for the Compustat 500, with labels marking ‘shortage’ years. The shaded areas denote ‘danger zones’ when Dominant Semiconductor Capital experienced differential *decumulation*. Figure 4 shows the percent deviation of Dominant Semiconductor Capital’s ROE from the Compustat 500 average, again, with marked years of ‘shortage’. These figures cover 1980 to 2021 – the years in which headlines about shortages were pulled from the *Nexis Uni* database.

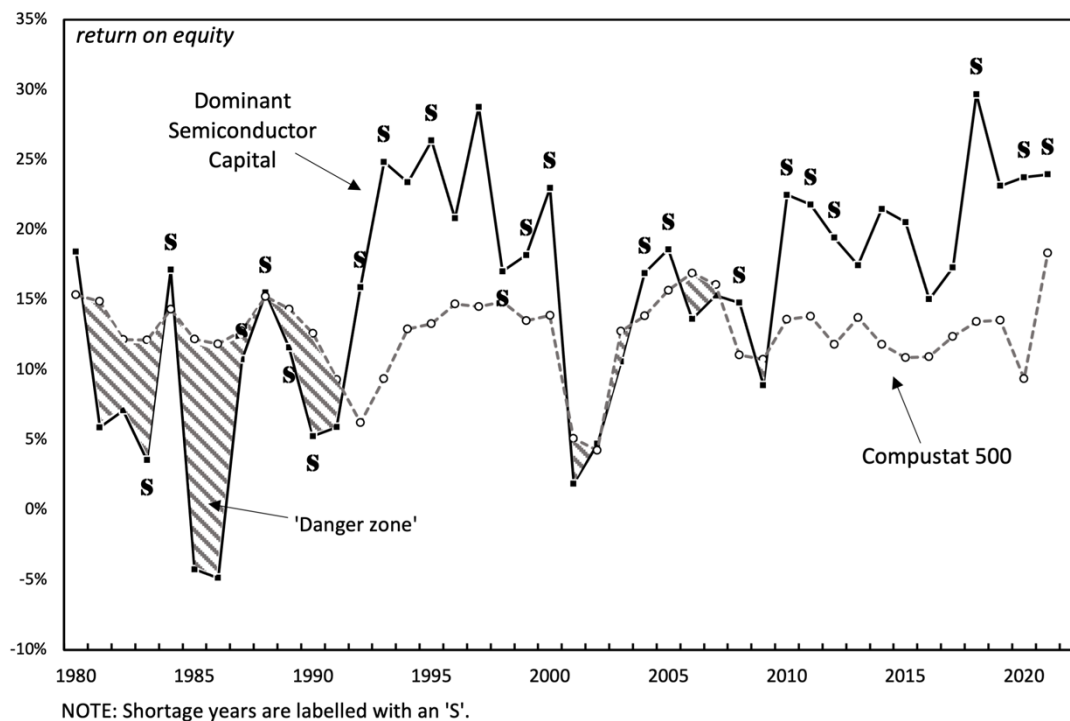


Figure 3. Return on equity and semiconductor shortages, Dominant Semiconductor Capital and Compustat 500 average (Method 1)

The pattern in Figure 3 is similar to the one found in Nitzan and Bichler (2002, Figures 5.7 and 5.8, 236-237). If we define a 'danger zone' as an uninterrupted period of a year or more in which Dominant Semiconductor Capital firms trail the average ROE, all but one danger zone between 1980-2021 ended with a perception of a 'shortage' the following year. In other words, the change in fortune from trailing to beating the average is almost always accompanied by an atmosphere of shortage. In addition, most years of differential accumulation are accompanied by 'shortage' (61%), in contrast to years of differential decumulation, in which 'shortages' are less common (28%).

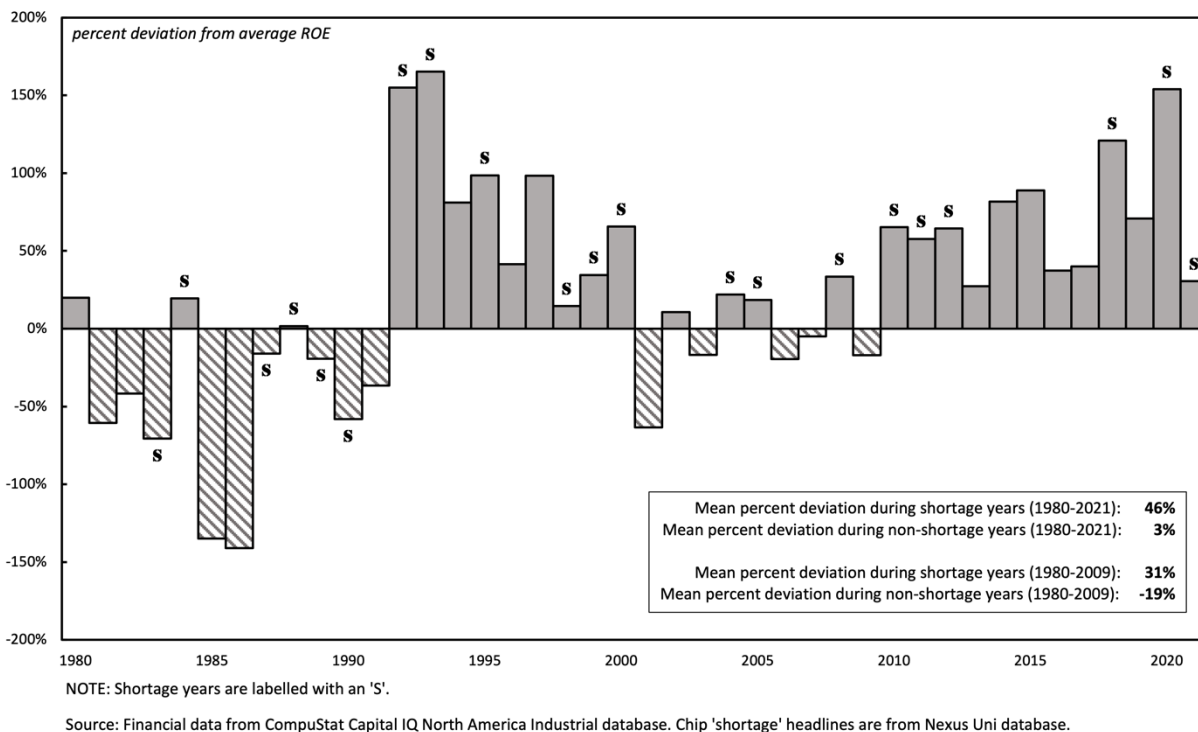


Figure 4. *Differential profitability and semiconductor 'shortages' (Method 1)*

In addition, as shown in Figure 4, there are substantial differences in differential profitability between shortage and non-shortage years. From 1980 to 2021, Dominant Semiconductor Capital exceeded the average by 46% during shortage years, while during non-shortage years, Dominant Semiconductor Capital only barely met the average. Between 1980 and 2009, Dominant Semiconductor Capital exceeded the average by 31% during 'shortage' years and trailed the average by -19% during non-shortage years. The 1980-2009 period is highlighted because, after 2009, there is evidence that the differential profitability of Dominant Semiconductor Capital was increasingly the result of a different strategy, that of mergers and acquisitions. This transition is examined in Section 5.3.

The evidence suggests a close relationship exists between the strategic limitation of chip production, differential profitability, and perceptions of a chip 'shortage'. On the one hand, it could be that a shortage is an *unintended* effect of 'over-shooting' the collective reductions in chip

production undertaken by Dominant Semiconductor Capital. On the other hand, the periodic regularity of shortages and their correlation with increases in profitability suggest an alternative: that the creation of the *perception* of shortages plays a crucial role in justifying the rise in chip prices from which firms differentially profit. As noted above, differential prices tend to generate resistance and reaction because they *redistribute* income. Framing the conscious reduction in investment and its consequent ‘shortages’ as an issue of supply and demand (i.e., outside any one firm’s control) obscures the role of power in this process. As Nitzan and Bichler point out:

The impact of differential accumulation through differential inflation depends crucially on the relative power of dominant capital. If this power is limited, raising prices faster than the average on the depth side is likely to be more than offset by lost market share on the breadth side, leading to net *decumulation*. (2002, 178).

If the strategy of differential accumulation through differential inflation appears to have been successful in the case of Dominant Semiconductor Capital, how did this group of firms achieve and maintain this relative power? The following two sections look first at the qualitative dimensions of this relative power (Section 5.2), and second, at how the accumulation strategies of Dominant Semiconductor Capital have evolved in response to centrifugal forces threatening to undermine that power (Section 5.3).

5.2 The 1986 US-Japan Semiconductor Trade Agreement

This sub-section details how, in the 1980s, a government-business coalition pressured Japan into coordinating a systematic reduction in semiconductor production. The purpose of this section is not to prove that shortages are always the result of government intervention or corporate collusion. Instead, it is to illustrate four qualitative aspects of capitalist power in the context of the semiconductor business. First, differential accumulation is a matter of the active application and augmentation of power and not simply production. Large semiconductor producers are not merely

passive receivers of ‘market-determined’ prices. They actively organize to restructure the global production of semiconductors to their differential advantage.

Second, the power of dominant capital *includes* the power of government institutions, which may act on behalf of or in defence of more general capitalist interests (e.g., in enforcing private property laws) or in the interests of an organized minority. In this case, the US government clearly acted in the minority interest of US merchant semiconductor firms, by taking unprecedented steps to secure the cooperation of the Japanese government and Japanese semiconductor producers to coordinate production and investment levels and reduce industrial capacity (Irwin 1996b, 11).

Third, Dominant Semiconductor Capital’s differential power does not necessarily imply a monopoly or cartel in the legal or conspiratorial sense. Setting aside the difficulties in proving such an argument, even the largest semiconductor firms do not hold a pure monopoly position, nor do they wield ‘absolute’ social power, nor does implicit cooperation necessarily imply outright collusion. However, the logic of differential accumulation compels firms to act in specific ways *if* they want to continue to accumulate differentially. This requirement means that dominant firms’ actions may be coordinated only to the extent that they each follow the same logic despite continuing to compete against one another. For instance, in this case, a main lasting effect of the US-Japan Semiconductor Trade Agreement was to change the strategy of the Japanese producers from internal breadth (through green-field investment) to one of external depth (through differential inflation).

Fourth, the exercise of power produces resistance and other unintended consequences that exceed the rationality of the agent’s intention. Differential depth, in particular, carries a high risk of generating new, oppositional power formations. In this case, the shortage of 1987-1990 allowed

Korean chip producers to gain market share and led to the creation of a new lobbying group for the US computer industry (Flamm 1996, 222; Irwin 1996a, 5).

5.2.1 The Semiconductor Industry Association and the creation of a business-government distributional coalition

While the popular narrative is one of US chip firms valiantly defending US chip production from government-backed Japanese firms, it was the US government that often “appeared to have been the cause and not just the effect of changes in the competitive conduct of Japanese semiconductor producers” (Flamm 1996, 127). Although the level of strategic cooperation of firms and governments to consolidate control over the production of chips reached a high-water mark in 1986, calls for government assistance on behalf of the US semiconductor began as early as the 1950s (127). As early as 1959, the Electronics Industries Association (EIA) “petitioned the Office of Civil and Defense Mobilization...to impose quotas on Japanese transistors” (128). The EIA at that time was split between electronics components producers and consumer electronics producers (129). On the one hand, components producers like semiconductor firms “favoured protection against consumer electronics imports from Japan—in which the bulk of inexpensive Japanese components entering the US market was embedded” (129). Consumer electronics producers, on the other hand, were worried about market access and were generally more deeply embedded in the Japanese market itself: some marketed Japanese products in the US for Japanese partners (Motorola and General Electric), some collected royalties from Japanese licensees (RCA), and others had significant ownership stakes in Japanese companies or controlled domestic Japanese subsidiaries (129). Thus, in 1968, the EIA ended up “testifying on both sides of trade policy issues” like protective tariffs (129).

By the 1980s, trade frictions between the US and Japan were compounded by the fact that, after embarking on a large-scale project of technological catch-up, Japanese semiconductor producers had surpassed US chipmakers in certain critical areas of production (Flamm 1996, 100). Central to this effort was the VLSI project, a series of multi-firm R&D subsidies organized by Japan's Ministry for International Trade and Industry (MITI) that "accounted for almost 40 percent... of Japan's national IC R&D effort in the late 1970s" (97-98). Out of the 22 significant results of the VLSI project, the majority were improvements in process technology that reduced costs and improved the quality of existing technologies (100). Consequently, Japanese chip producers could vastly improve chip yield (the number of working chips produced in a single silicon wafer), allowing them to make chips at a much lower cost than their American competitors, particularly in DRAM, or dynamic random-access memory, manufacturing.

The creation of the Semiconductor Industry Association (SIA)—the new "lobbying arm" of large US semiconductor firms—in 1977 was a direct consequence of the growing awareness by US chip firms of the significant gains Japanese producers were making in international markets (Flamm 1996, 138). As Irwin notes, the SIA reflected the common interests between otherwise openly antagonistic firms:

Each of these firms competed fiercely with one another on certain dimensions—suing each other over alleged patent violations, for example, or even conducting espionage against one another—but they could agree on several common policy objectives, such as obtaining greater patent protection for chip designs, improving the tax treatment of R&D investment, and heightening political awareness of the emerging Japanese competition. (Irwin 1996b, 21)

The organization successfully convinced the US Senate to "order the US International Trade Commission (ITC) to launch an informational investigation into the competitive position of the US semiconductor industry" by arguing that the Japanese were poised to steal market share from

US producers through illegal trade protections (139). While evidence of illegality was thin, generous political donations made by the four largest chip firms during this period no doubt added gravity to the urgency of their claims (Irwin 1996b, 21). Irwin notes that “disbursements of these PACs [political action committees] appear to be related to the trade dispute with Japan: payments totalled \$354,318 at the peak of the dispute in 1985-86, 40 percent higher than in 1983-84 and 17 percent higher than in 1987-88 after trade tensions had simmered down” (1996b, 22).

By the end of 1979, during a 16K DRAM ‘shortage’ in the US, Japan’s three largest chip producers had managed to capture 40 percent of the US chip market (Flamm 1996, 139-140). The SIA argued that the Japanese had achieved this capture by ‘dumping’ using a two-tiered pricing structure: keeping prices high in the Japanese domestic market and exporting chips ‘below cost’ or ‘below fair value’ to gain market share in foreign markets (141). The evidence for this accusation, Flamm notes, was scarce and primarily anecdotal (138). Moreover, it is unclear how such claims could have been realistically investigated—not only due to the complexities of the different US and Japanese domestic market structures but also due to the inherent subjectivity of determining what constitutes ‘fair value’.

Despite the lack of evidence, US chip producers soon began to complain that “the Japanese were selling below the cost of production in *both* [US and Japanese] markets” (Flamm 1996, 142). One of the most curious charges was ‘quality dumping’ (145). Japanese firms had successfully improved their chip manufacturing standards to the point that US chip customers reported defect rates of Japanese DRAMs at “one-half to one-third those experienced with comparable American products” (145). Instead of denying the truth of these reports, US chip producers “charged that sales of higher quality Japanese products at the same price as American products reflected a form of ‘dumping’” (145). At the same time, Texas Instruments (TI), the only US producer who had

access to the Japanese chip market, complained in the early 1980s that the SIA's complaints had caused Japan to voluntarily restrain its exports, lowering prices in Japan (and hurting TI's profits) (143). Fearing anti-dumping actions by the US Trade Representative (USTR), Japanese producers had suspended sales to the US spot market for DRAMs and announced plans to open US production facilities "in a complementary bid to reduce trade frictions" (143).²⁰

5.2.2 Pressuring Japan

Political pressure on Japanese producers to reduce exports to the US and raise prices increased steadily through the early 1980s. In 1981, Japanese producers introduced new 64K DRAMS ahead of U.S producers and "by early 1982...the Japanese share of the US 64K DRAM market [stood] as about 70 percent" (Flamm 1996, 148-149). Again, US producers pressed the Commerce Department "to investigate charges that the Japanese were selling 64K DRAMS below 'fair value'" (149). In response, Washington warned the Japanese government that the "Commerce department might begin to 'monitor' Japanese import prices" (149). However, as the investigation was set to begin, "DRAM prices suddenly doubled, Japanese suppliers began rationing US customers, and it was reported that Japanese companies were cutting back US exports in order to blunt moves toward trade restrictions on DRAM imports" (149). Companies confirmed to reporters that "they were reducing US exports to alleviate trade friction" at the behest of the Japanese (and US) governments (149-150). In 1983, another 'shortage' appeared within a year after introducing export reductions, and chip prices shot up (151). Ironically, following the 1983 price increase, US producers immediately began accusing the Japanese producers of collusion and price-fixing, and the US Justice Department opened an antitrust investigation into 'excessively high prices' (152).

²⁰ A spot market is a more informal market for chips based on short-term contracts and populated by trading firms, manufacturing firms, and customers.

While the case eventually faded away, the mixed and contradictory accusations reportedly left Japanese producers somewhat confused (151). When the 1983-1984 'shortage' subsided, criticism of the Japanese returned with force. Antidumping petitions were filed against three types of chips, while a private antitrust suit and several complaints of unfair trade practices (under the infamous Section 301 of the 1974 trade act) were filed in 1985 (160). All but the antitrust case was suspended with the negotiation and signing of the Semiconductor Trade Agreement (STA) in September of 1986 (160).

5.2.3 The 'shortage' of 1987-1990

The 1986 STA established export price controls for several kinds of semiconductors and called for extensive monitoring of DRAM and EPROM prices by the Japanese government and a host of other products. The agreement stipulated that the Japanese government would "take appropriate actions available under law and regulations in Japan, including ETC [the Export Trade Control ordinance], in order to prevent dumping," and implementation moved quickly (Flamm 1996, 177). MITI suspended new export licences for 256K DRAMs and advised Japanese producers to set export prices at or above 'fair value', which was dictated by the US Commerce Department (177). It also set up a Forecast Committee, "whose task it was not only to publish forecasts but also to help in 'correction of imbalances' between Japanese supply and demand for products covered by the price monitoring framework" (177-178).

At first, producers found "creative mechanisms to evade the newly minted MITI controls," and many firms did not cut back production to the levels encouraged by the Forecast Committee (Flamm 1996, 179). In particular, TI Japan and NEC, the largest producers of 256K DRAMs, were "reluctant to follow MITI's new 'guidance' on production and export volumes" (180). However,

it is possible that a turning point occurred in March 1987, when President Ronald Reagan declared that “prohibitive (100 percent) tariffs would be imposed on \$300 million worth of imports from Japan... [the] largest and the first [unilateral retaliatory trade] action against this US ally in the postwar period” (Irwin 1996b, 11). As Irwin notes, “the sanctions were crafted to hit the exports of the principal Japanese semiconductor producers—such as NEC, Toshiba, Hitachi, and Matsushita—but not entail significant consumer losses” (53). More than anything, these measures may have compelled Japanese chip producers to follow the ‘guidelines’ set by MITI.

Eventually, producers began to comply with the pressure. For instance, TI Japan announced it would “slash its output of 256K DRAMs by 13 percent to comply with MITI’s wishes,” and NEC announced it would cut production by 40 percent (183-4). By the spring of 1987, “both the production and export of DRAMs by Japanese companies had been placed under fairly tight MITI controls,” and prices began to rise worldwide (185). As Flamm notes, “annual rates of change in price hit all-time historical highs in 1987 and 1988, after the STA went into effect,” and in a business defined by constant falling prices, *positive* changes were recorded for the first time for memory chips in 1988 (237). Flamm estimates that the “‘guidance’ supplied to producers in restraining investment levels continued at least through early 1988, and therefore probably affected supply through at least 1989” (272).

5.2.4 The privatization of price controls

By early 1988, rising prices brought complaints by chip users, as well as fears of another ‘shortage’ – one that was beginning to be perceived in both the US and Europe (Flamm 1996, 192). Even as prices continued to climb, however, “MITI’s control framework was extended into new areas” (192). For its part, “the increasing signs of shortage satisfied the American government that

MITI had acted forcefully to increase chip prices in worldwide markets” (193). However, to counter public pressure from the chip users, who were complaining about the high prices, the US again “switched its public posture to one of encouraging MITI not to restrict chip production” (193). In truth, however, the government was “less than unequivocal...officials were not eager to see Japanese firms increasing their chip capacity to meet the looming shortage, preferring to see American companies ‘re-enter’ the DRAM market” (194). Far from acting to remove limits, “American trade negotiators continued to press MITI to limit investments by Japanese firms in new capacity well in 1988” (195).

Due to increasing complaints by Japanese chip users, “MITI became considerably more reluctant to spell out the precise nature of its actions in public” and began to advocate for firms to “extricate themselves from the inclination toward excessive competition” rather than relying on the guidance of MITI (Flamm 1996, 184-185). The idea that it was “indispensable for manufacturers to make their own efforts...to establish prices in accordance with the balance of supply and demand” was increasingly promoted both by private firms and the government (186). For instance, one 1987 MITI report “called for a considerable amount of coordination among rival firms in the semiconductor industry” and “specifically called on semiconductor producers to cooperate in planning investments and in matching production to forecast levels of supply and demand” (187). In effect, both government and businesses agreed to ‘privatize’ the price and production controls implemented through the framework of the STA.

5.2.5 The uncanny accuracy of production ‘forecasts’

Perhaps the most obvious evidence that MITI’s production forecasts acted as semi-compulsory guidelines was their remarkable accuracy during the 1987-1989 ‘shortage’ years

(Flamm 1996, 201). For instance, between the second and fourth quarter of 1987, “forecasts issued for 256K, and 1M DRAM production levels three and six months out typically fell within 10 percent of actual output” (197). This level of accuracy continued even after “the invisible hand, rather than the government’s, ostensibly ruled the market,” attesting to the success of the privatization of price controls (197). Some analysts argued that the shortages were due to other factors than the purposeful restriction of production toward the aim of price increases (197). One prominent argument was that “unexpected yield problems were a major factor in the shortage then developing [in 1988]” (197). However, Flamm points out that “if such yield problems, in the aggregate, played a significant role in creating shortages...it is hard to see how they could have been unexpected, given the accuracy of the MITI forecasts” (200). “A continuous history of coming within 10 percent of three-and six-month production forecasts,” he reasons, “in a product requiring over two months of processing on the production line, suggests that unanticipated yield problems could not have been a major issue” (200).

5.2.6 The entrance of Samsung and the second STA

The ‘shortage’ also produced centrifugal forces: principally, it allowed emerging South Korean chip producer Samsung, who was “not subject to the political and legal pressures faced by Japanese chip-makers,” to undercut DRAM prices and vastly expand their market share. By the time the shortage atmosphere had dissipated, Samsung had become “the largest producer of 1M DRAMs in the world” (Flamm 1996, 222). On the other hand, Samsung’s successful expansion incurred the ire of its competitors, and in 1990 Micron Technologies (one of two remaining US DRAM producers) “raised the possibility of a dumping suit against Korean vendors” (226). Like the Japanese, Samsung quickly obliged and cut DRAM production by 20% (226). In addition,

“dumping cases against Korean exports of DRAMs were filed in the European Community (EC) in 1991 and in the US in 1992” (224). “Faced with stiff antidumping duties,” writes Irwin, “the Korean industry and government proposed in January 1993 a bilateral semiconductor trade agreement fashioned on the earlier one with Japan,” in which “the Korean industry promised to monitor prices of export sales to the United States” and ensure “demonstrable and measurable results in terms of increasing sales in Korea of US semiconductors and semiconductor equipment” (Irwin 1996b, 60).

In June 1991, a second STA was signed, replacing the one set to expire the following month (Flamm 1996, 223). This agreement removed the ‘fair market value’ pricing floors, introduced a fast-track antidumping procedure, and retained the extensive price monitoring system. Flamm argues that practically speaking, little changed (224). MITI continued to collect industry data and publish ‘forecasts,’ and “thus a variety of well-established mechanisms designed to constrain pricing in world semiconductor markets continued in their original or revised form” (224).

5.2.7 Summing up

What can be concluded from these events? First, the events of the 1980s show that the notion of ‘shortages’ is unexplainable without reference to the broader *power* dynamics of the sector. Competition from Japanese producers was countered by coordinated political pressure to reduce competition and restrain production (Flamm 1996, 206). While coerced initially, Japanese firms quickly understood that they had as much to gain from higher margins as US producers. As one Japanese executive put it: “since the Semiconductor agreement, we [Japanese DRAM makers] have moved from competing for market share to market sharing” (quoted in Flamm

1996, 215). The profits from differential inflation could only have been achieved through a coordinated effort to curtail production.

Second, the emergence of the SIA and the success of its lobbying efforts marks the emergence of a loose ‘distributional coalition’ (to use Mancur Olson’s term) of semiconductor firms and government organs that coordinated the limitation of production to increase chip prices. If the correlation between ‘shortages’ and the differential profitability of Dominant Semiconductor Capital is any indication, the development of this coalition was far more critical than any initial profits gained from the DRAM ‘shortage’ of 1987-1990. Irwin notes, “only two US merchant firms (TI and Micron) remained in the DRAM market to benefit from the antidumping actions. DRAM sales reportedly accounted for as much as 60 percent of TI’S profits in 1988, and Micron’s sales rose by a factor of six between 1986 and 1988” (1996b, 11-12). On the whole, the creation of this ongoing government-business coalition was much more valuable than any initial monetary gains of these two firms. As Irwin argues, “such a sectoral agreement [as the STA] is attractive from the perspective of virtually any import-competing industry because it virtually guarantees the institutionalization of trade policy for that industry” (11). Moreover, “once the agreement was in place, it required monitoring and at some point, renewal, or renegotiation... [providing] a natural rationale for ongoing contacts between the industry and the government, providing the industry with easy access to key policymakers and allowing close industry-government ties to develop” (12).

Third, compounding the cooperation of governments and sanctioned by them is the open exchange of price and production information, for example, through the ‘demand forecasts’ published periodically by MITI and by organizations like the SIA. Publishing these forecasts allows producers to adjust new capacity investments to avoid unprofitable ‘overproduction’

without officially breaking the law. In addition, these structures of dominant power tend to recede into the background of the public consciousness as the political contention engendered initially by the creation of such a coalition is replaced by the matter-of-fact designation of the semiconductor business as a ‘strategically important’ industry.²¹

A fourth conclusion is that the relative power of dominant capital is never total but is constantly shifting. An unintended effect of the ‘shortage’ of 1987-1990 was the entrance of new competitors—namely Samsung—highlighting the centrifugal forces that are generated and exceed efforts by dominant firms to differentially accumulate through the systematic limitation of production. In the next section, these centrifugal forces and their effects on the semiconductor business are examined in greater detail.

5.3 A new breadth millennium

In their 2002 book, *The Global Political Economy of Israel*, Nitzan and Bichler track the global shift from depth to breadth starting in the late 1980s (294). This transition, they argue, rested on three “breadth-related poles”: *capital decontrols*—referring to the increasing opening up of national markets to foreign investment and ownership, particularly under free trade agreements like NAFTA and the EU—*privatization*, and expansion into ‘*emerging markets*’ (295). While each of these certainly affected the shift to breadth within the semiconductor business, this section focuses on a fourth set of factors—the centrifugal forces within the semiconductor manufacturing business—and their relation to Dominant Semiconductor Capital’s shift to a breadth strategy in the mid-1990s. It argues that centrifugal forces at work in the semiconductor business led to a steady acceleration in the growth of new firms between approximately 1985-2005. This growth

²¹ For a discussion of whether or not the semiconductor business is ‘strategic’, see Flamm 1996, Ch. 7.

first destabilized the differential power of dominant semiconductor firms, leading to increased uncertainty about their relative power. Subsequently, the rapid growth of new firms became the basis for a new breadth period, as large firms vastly increased the pace of corporate amalgamation, stabilizing their relative power and vastly increasing their differential earnings in the process.

5.3.1 Centrifugal forces

As mentioned in Section 2.3.3, mergers and acquisitions are reliable forms of differential accumulation. As Nitzan and Bichler note, merger and acquisition activity “kill[s] three birds with one stone: it directly increases differential breadth; it indirectly helps to protect and possibly boost differential depth (relative pricing power); and it reduces differential risk” (2009, 330). In addition, the *integrative* nature of information technology—reliant upon interoperability, the growth of networks, and its capacity to link social activity across geographic space instantaneously—lends itself to mergers and acquisitions (Nitzan and Bichler 2002, 294). Due to this integrative tendency, the expansion of the ‘network’ constantly introduces “new players, new forces, and new rules,” which “if left unattended, tend to destabilize established power and undermine profit” (294). Nitzan and Bichler argue that “the most common way of containing these centrifugal forces is through the centripetal, counter-force of corporate amalgamation” (294-295). Thus, a breadth strategy may be more critical in a rapidly changing and expanding business landscape than in one which is more predictable, growing slower, and less dynamic. If this is the case, one would expect the semiconductor business to be an early adopter of a breadth strategy of mergers and acquisitions. However, it was not until the mid-1990s that acquisition spending rose to close to the corporate average (see Figure 5 and Figure 8). Why did it take so long for the semiconductor business to adopt it as a strategy?

One possible reason is that there were simply too few firms to take over. Before 1982, less than 20 semiconductor manufacturing firms were listed in the Compustat North America database. High start-up costs and a small, mostly government market also meant that mergers and acquisitions, when they did happen, were minor in relation to other expenditures. In addition, non-US-based chip producers tended to be large, vertically intergraded consumer electronic firms and government-backed ‘national champions’ – meaning they were too large, legally protected or both. These firms were neither available nor vulnerable to acquisition until the 2000s, when national economies ‘globalized’, capital controls were deregulated, and many large conglomerates ‘spun off’ their semiconductor operations into separate businesses (McClellan 2011, 2-3).

However, despite the initial high concentration of firms, as computing grew in importance to the industry as a whole, and as semiconductor firms continued to experience the appearance of periodic shortages, important centrifugal forces generated more and more new firms. For one, the frequent appearance of shortages, along with higher prices, presented opportunities for new competitors to enter and steal market share by undercutting the other large firms (discussed in Section 5.2).

Second, high prices and concentrated power are strong incentives for institutional chip customers to encourage greater fragmentation in the sector, and enterprising computing firms like Apple often actively seek to undermine the power of large chip producers like Intel. For instance, in 1987, Apple worked with firms VLSI and Acorn to create a jointly owned, independent company to design the chip for its Newton personal digital assistant (Nenni and McLellan, 2013, 35). While the Newton was not a commercially successful product, the joint venture, ARM, became one of the largest suppliers of semiconductor IP and the standard microprocessor architecture in mobile phones (170-171).

Third, the speeds at which semiconductor technology, and information technology in general, were developing, presented new opportunities for leading-edge or highly specialized producers to capture market share as new industrial uses emerged. For instance, as computing power increased and became cheaper and smaller, the range of services to which computation could be put grew. Despite these opportunities, large firms like Texas Instruments and Intel, which in the 1980s controlled every aspect of the chip-making process from design and manufacturing to packaging and testing, tended to focus on mass production of only the most profitable products. The growing gap between these firms' product lines and the diversity of industrial needs opened up opportunities for smaller companies to specialize in these riskier and less profitable markets. A result of this increasing diversity of custom chips was the emergence of 'fabless' semiconductor firms, who design but do not manufacture their chips. The growth of fabless firms, which initially operated by renting excess capacity from existing manufacturers, only accelerated with the introduction of contract manufacturers called 'foundries'. Semiconductor foundries do not design their chips but devote their operations to producing chips designed by others. The most important of these is the Taiwan Semiconductor Manufacturing Company (TSMC), founded as a joint venture between the Taiwanese government and Dutch electronics manufacturer Philips—another example of a chip-user acting as a centrifugal force—and other private investors (Nenni and McLellan 2013, 76). TSMC has become one of the semiconductor business's largest and most powerful firms (83).

In short, many strong centrifugal forces threaten to destabilize the differential power of dominant semiconductor firms.²² The following analysis argues that the acceleration of these forces in the 1990s destabilized the power of Dominant Semiconductor Capital, giving rise to the

²² These forces also worked in parallel and combination. For instance, Apple contracted eSilicon, the first "fabless ASIC" firm, to supply the system chip for the original iPod (41).

need for a corresponding strategy of corporate containment through mergers and acquisitions. Here, again, the rapid pace of technological change presented a *problem*, rather than an aid to the differential accumulation of dominant firms, and the solution was to subjugate technical change by buying up—rather than competing with—new firms.

5.3.2 Quantitative analysis

The following quantitative analysis uses the growth of new firms as a quantitative proxy for these centrifugal forces, in order to show their two-fold impact. First, the rise in new firms through the mid-90s and early 2000s undermined certainty about the relative power of Dominant Semiconductor Capital, quantitatively expressed as greater volatility in capitalization. Second, the new firm growth created a pool of new targets for acquisition, and an internal breadth wave of mergers and acquisitions began around the same time. At first, the pace of new firms' emergence ran ahead of dominant firms' mergers and acquisitions activity. By the 2010s, the trend reversed, and the increasing pace and size of M&A activity significantly reduced the overall number of firms. The result was a massive increase in the differential earnings of Dominant Semiconductor Capital and a reduction in the volatility of Dominant Semiconductor Capital's capitalization.

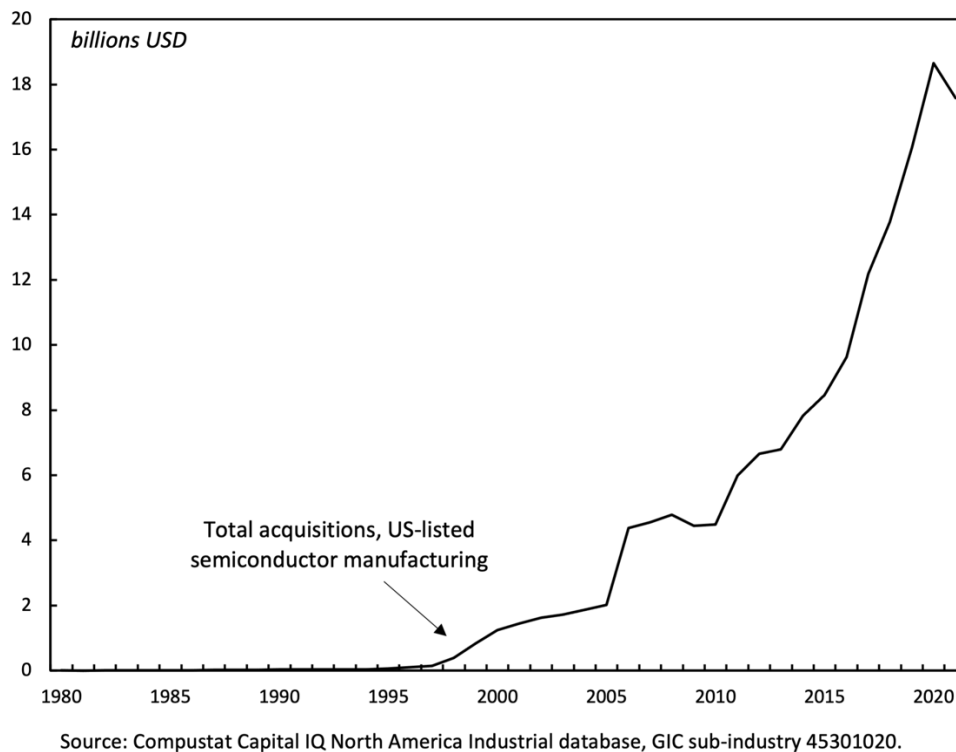
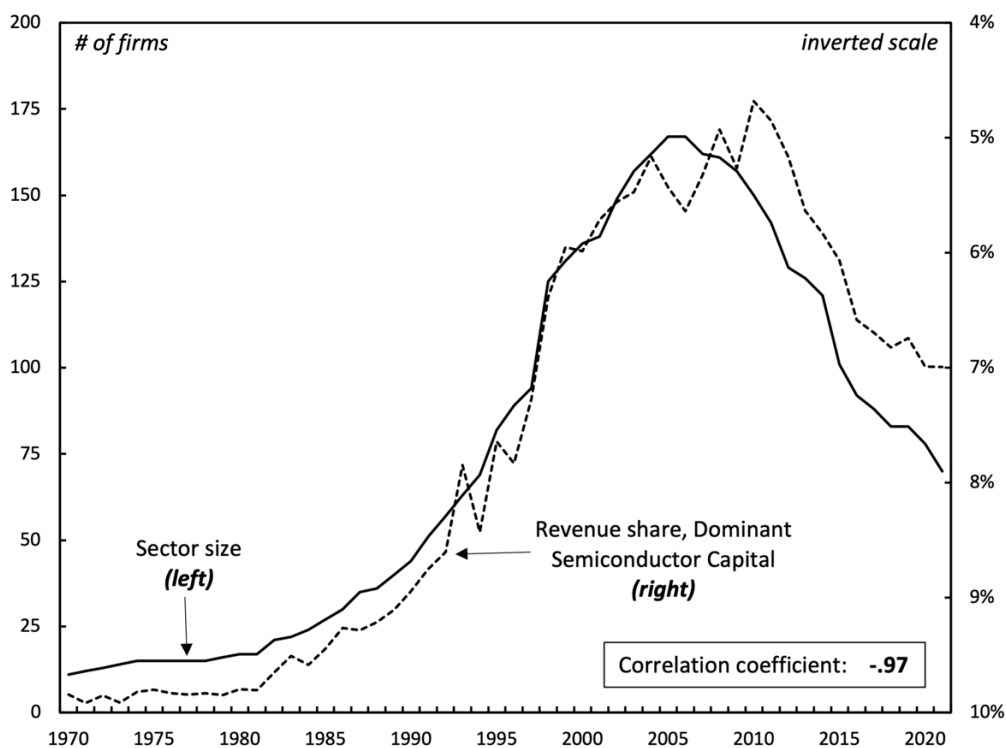


Figure 5. Annual volume of acquisitions by monetary value, semiconductor manufacturing

Figure 5 illustrates how dramatic this ‘breath’ wave was. Acquisitions were negligible during the 1970s and 1980s but, starting in the mid-1990s, rose to several billion dollars a year. Though M&A activity slowed somewhat in the aftermath of the dot-com bust and again during the financial crisis of 2008-09, the sector has followed a path of more or less exponential growth in mergers and acquisitions activity.

5.3.2.1 ‘Industrial’ growth and uncertainty

The rise of new firms is, in effect, green-field investment – by increasing the size of the overall earnings pie, it undermines the differential control over production held by Dominant Semiconductor Capital. This can be shown quantitatively in multiple ways.



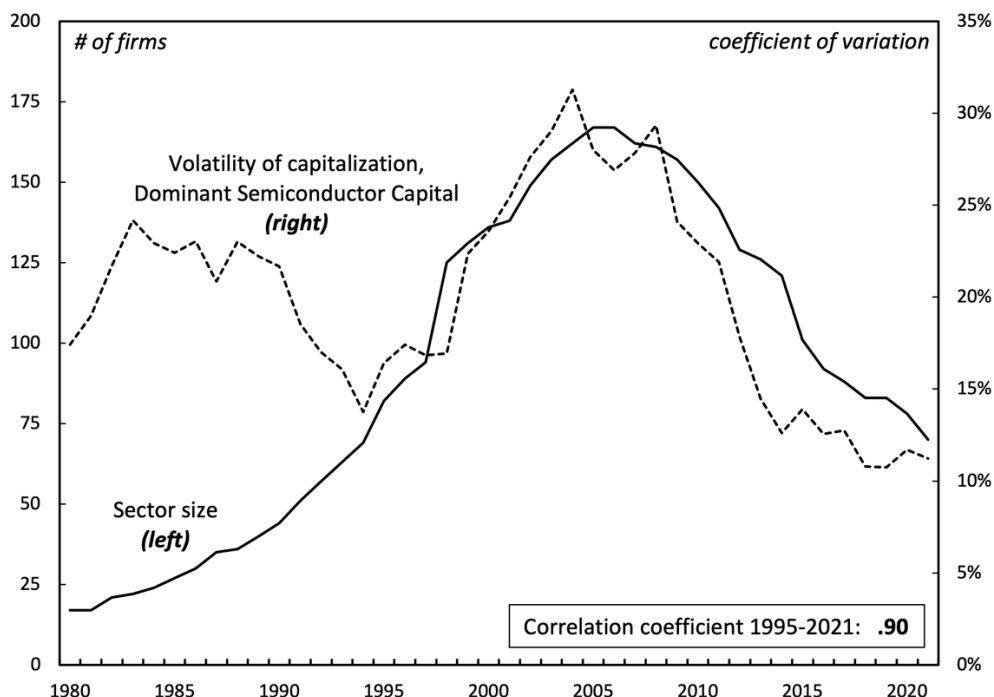
NOTE: Sector size is calculated as the number of firms listed in the Compustat database. Dominant Semiconductor Capital is calculated as the top ten firms by market capitalization. Correlation is calculated using the log of dominant capital's share of revenue.

Source: Compustat Capital IQ North America Industrial database, GIC sub-industry 45301020.

Figure 6. Sector size and Dominant Semiconductor Capital's share of the revenue (Method 1)

Figure 6 shows the sector size, measured as the number of semiconductor manufacturing firms listed in the Compustat database and the share of total revenue for the sector held by Dominant Semiconductor Capital. As the number of firms increases, the percentage of total sales captured by Dominant Semiconductor Capital shrinks. The tight negative correlation between the two series (-.97, note the inverted right scale) suggests that the sector's overall growth undermined the share of revenue of Dominant Semiconductor Capital, threatening their relative power. It also illustrates the danger for dominant capital of 'too much' green-field investment. As total production expands, Dominant Semiconductor Capital must grow *faster* than the average to maintain the same revenue share. However, if retaining a differential profit *margin* is reliant on

tightly restraining the expansion of production, any increase in production to keep up with the sector's growth will tend to work against this strategy.



NOTE: The size of the sector is measured as the number of listed firms in the Compustat database. Dominant Semiconductor Capital is measured as the top ten firms by market value. The volatility index is constructed by calculating the coefficient of variation (the standard deviation from the mean) of the rate of change in the average capitalization of the top ten semiconductor firms. The volatility index is shown as a 10 year trailing average.

Source: Compustat Capital IQ North America Industrial Database.

Figure 7. Sector size and capitalization volatility, Dominant Semiconductor Capital (Method 1)

Figure 7 confirms the destabilizing effect of new firm growth on the power of Dominant Semiconductor Capital using another way of measuring power: capitalization. According to capital as power theory, capitalization represents a symbolic estimation of a firm's ability to generate risk-adjusted earnings in the future. Capitalization is forward-looking and, in principle, a universal process. Nitzan and Bichler write: "if it generates earning expectations, it must have a price, and the algorithm that gives future earnings a price is capitalization" (Nitzan and Bichler 2009, 158). Because the future is inherently uncertain, and since capitalization is a future-oriented process, the

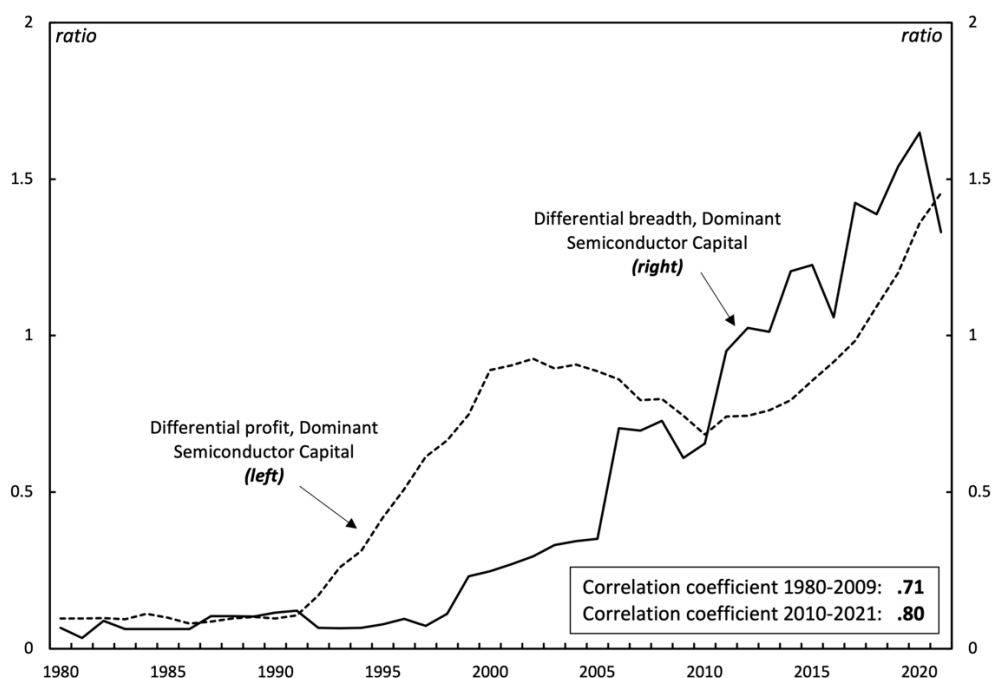
volatility of a firm's capitalization can be understood as a quantitative proxy for measuring capitalists' level of certainty or uncertainty about their estimations of the future. In short, all else remaining the same, the less *certain* capitalists are of those future expectations, the more change or volatility one might see in a firm's capitalization.

Figure 7 shows the relationship between the size of the sector, measured as the total number of listed firms in the Compustat database, and the volatility of Dominant Semiconductor Capital's capitalization. Volatility here is calculated as the coefficient of variation of the rate of change in the average market value of Dominant Semiconductor Capital.²³ The correlation between new firm growth and the volatility of Dominant Semiconductor Capital's capitalization between 1995-2021 (.90) suggests that as the sector expanded through the 1990s, the power of Dominant Semiconductor Capital became more *uncertain*. This pattern makes intuitive sense, as the growth of the number of firms is simply the growth in the number of potential challengers or competitors. With each new firm, the future power of current dominant firms is potentially diminished by an unknown degree. Therefore, a larger field of play is inherently more dynamic and riskier. The rise in volatility suggests that growth in the sector outpaced dominant firms' ability to buy up competitors, increasing uncertainty around Dominant Semiconductor Capital's differential power. By 2010, the trend had reversed: as the number of firms in the sector decreased and M&A activity skyrocketed, Dominant Semiconductor Capital's capitalization (read: estimated relative future earning capacity) became less volatile.

²³ The coefficient of variation is measured as the standard deviation from the sample's mean.

5.3.2.2 Differential buy-to-build

Figure 8 compares the differential buy-to-build ratio and differential profitability of Dominant Semiconductor Capital. As noted in Section 3.4, differential buy-to-build represents the emphasis of dominant semiconductor firms on investment in M&A over the expansion of productive capacity, relative to a benchmark. Differential profitability is measured here as a ratio of the average net income of Dominant Semiconductor Capital to the average net income of a Compustat 500 firm.



NOTE: Differential profit of Dominant Semiconductor Capital is calculated as the average annual net income of the top ten semiconductor manufacturing firms by market capitalization, divided by the average net income for the Compustat 500. Differential breadth is calculated by dividing the ratio of annual acquisitions to annual capital expenditures for the top ten firms by the same ratio for the Compustat 500. Series are shown as ten year trailing averages.

Source: Compustat Capital IQ North America Industrial database.

Figure 8. Differential buy-to-build and differential profitability, Dominant Semiconductor Capital (Method 1)

Overall, there is a strong correlation between differential buy-to-build and differential profitability within Dominant Semiconductor Capital, particularly during 2010-2021 (.80). Up

until the mid-1990s, both differential buy-to-build and differential profitability remained fairly level, possibly because of the lack of takeover targets in the semiconductor business. Starting from the end of the 1990s, differential buy-to-build rose steadily, while differential profitability initially rose, then moved sideways/fell through the 2000s. After 2010, both differential profitability and differential buy-to-build rose sharply, suggesting that the rapid gains of Dominant Semiconductor Capital during this period were caused at least in part by their differential pace of acquisitions.

If the creation of new ideas represents one pole of technological change within the semiconductor business, then the physical manufacture of the products of those ideas represents the other. On the one hand, the rapid pace of change in semiconductor development, combined with the periodic differential *decumulation* of Dominant Semiconductor Capital, suggests a certain degree of failure to profitably limit the creation and diffusion of new ideas. On the other hand, the regular appearance of ‘shortages’ suggests that dominant semiconductor firms have primarily relied on the strategic limitation of the *physical manufacture* of chips for differential accumulation. The evidence in this section indicates that this strategy may have shifted as new design, IP, and other fabless firms emerged as targets for mergers and acquisitions. It also confirms the capital as power thesis that, under certain conditions, M&A activity is a reliable way to accumulate differentially. Finally, it shows how, as the new growth of semiconductor firms since the mid-1990s created a wealth of takeover targets, dominant firms adapted to the centrifugal forces threatening their differential power by shifting accumulation strategies.

5.3.3 A return to depth?

The current trade conflict with China parallels the conflict with Japan in the 1980s, not least because the semiconductor business appears to be implicated at the highest levels. While it

is still a few ‘production nodes’ behind US and Taiwanese semiconductor firms, The Chinese ruling elite is pouring resources into achieving technological parity, if not supremacy, with the other technological superpowers (Capri 2020, 30). China is already the largest purchaser of semiconductors globally, and its technology sector is highly integrated with foreign capital (23). The combination of policies designed to encourage foreign investment and the desire to gain access to China’s rapidly growing technology sectors have fueled a steady migration of semiconductor business into the country (Chu 2013, 188). A 2020 Hinrich Foundation report lists 22 semiconductor firms’ joint ventures with China between 2014 and 2018, while “in China’s computer sector, foreign-invested enterprises accounted for 59 percent of industry assets and 57 percent of industry profits in the manufacture of computers” in 2013 (Capri 2020, 34).

At the same time, the unleashing of industry under China’s industrialization policy has also sparked fears of ‘overproduction’. For instance, Capri estimates that “the possibility of an over-supply of NAND and DRAM chips would seem likely, at some point, which would drive down global market prices” (2020, 29). In other words, China appears to be pursuing a similar strategy as Japan did in the 1980s – a strategy of green-field investment that some worry will undermine the differential profitability of other firms. As Capri notes, “none of this bodes well for the world’s existing players” (29).

It is possible that the trade war is making it increasingly difficult to continue the integration of Chinese firms into an increasingly global Dominant Semiconductor Capital. Starting in 2016, the Committee on Foreign Investment in the US (CFIUS) began closely scrutinizing M&A activity between the US and China and shut down several high-profile cases (46-48). Most notably, former President Donald Trump blocked a \$117 billion acquisition of semiconductor manufacturing firm Qualcomm by (at the time) Singapore-based Broadcom – a deal which would have been one of the

largest in history. Does this point to an end to the breadth wave of the 2010s? Since 2020, the world has been gripped by a global chip ‘shortage’; in turn, chipmakers have booked record profits. Whatever the future holds for the US and China, Dominant Semiconductor Capital will likely continue to play a starring role.

Section 6. Conclusion

Semiconductor technology and semiconductor firms do not exist in a vacuum – they are enmeshed in complex ways with social forces outside the control of any dominant group. The interests of the semiconductor business cannot wholly escape the influence of government and military agencies, computer and consumer electronics producers, or the tastes and fashions of computing culture at large. Neither are these institutions, individuals and dynamics easily separable from the sphere of semiconductor ‘business’ interests. From a capital as power perspective, the stock price of publicly traded semiconductor firms capitalizes the power of global intellectual property rights, Chinese government policy, innovations in the computing industry, worker migration and everything else that may bear on the future profitability of the company. Each of these is, in turn, shaped by the logic of differential accumulation.

Given the porous, shifting nature of social power (however formally and ritualistically capitalized), what can be said about the dynamics of technology and capitalism based on the evidence presented here? First, while ‘business’ profits through the strategic sabotage of ‘industry,’ this sabotage appears to be modulated by conditions often outside the control of any one firm or group of firms. In the information technology sector, one of those conditions may be the technology itself. Information technology is not simply an inert product with a relatively stable set of uses, like oil or wheat, but primarily a body of ideas and specialized knowledge. In a sense, what defines ‘technology’ in this context is this open-ended frame of reference: it presupposes its

own evolutionary process. If technology is in this sense defined as one leading edge of the creative processes of a community, then the business of ‘technology’ carries with it an inherent expectation of less than total control. As such, the business of technology represents both a kind of outer limit of the possibilities of (present) control and, simultaneously, a limitless imaginary of new *forms* of (future) control.

To use the terminology of capital as power, this liminal space between the current state of technology and the promise of new advances holds a strong ideological potential for both extreme ‘hype’ and extreme ‘risk’.²⁴ These terms are ideological because they refer not to human-oriented goals but the potentialities of *power*. This technological imaginary of power may play a central role in the ideological power of capitalism by shaping and limiting our ability to envision possible futures to its terms. The importance of semiconductor technology in techno-capitalist visions of the future—with their assumptions of a virtually limitless, exponential growth in computing power—may in part lie at the heart of the integrated circuits’ seemingly uncontroversial status as a ‘strategic’ technology, and in the continuing power large semiconductor firms hold over society at large.

6.1 Mumford’s warning

That said, it is worth returning to Lewis Mumford’s warning about equating all technological advancement with the development of human potential and well-being. His warning was this:

²⁴ See Nitzan and Bichler 2009, Ch, 11 for a discussion of these ‘elementary particles’ of capital.

While any new technical device may increase the range of human freedom, it does so only if the human beneficiaries are at liberty to accept it, modify it, or reject it: to use it where and when and how it suits their own purposes, in quantities that conform to those purposes. (Mumford 1971, 185).

It is not difficult to find examples of how computing technology has failed us in these respects. To name a few: the military use of semiconductors in ever more sophisticated and lethal weapons systems has undoubtedly made the world a more dangerous place; cheap computing power has led to the rise of socially and ecologically destructive practices like digital addiction and cryptocurrency mining; while the ownership of mobile phones—devices designed for the widespread and invasive surveillance of populations—is now a precondition for obtaining even the most basic social necessities of life. In this context, in the words of Mumford, the pressure to “forego all modes of activity except those that call for the unremitting use of the ‘machine’ or its products” is perhaps itself a form of social sabotage (329). Thus, the conclusion of this paper is *not* that technological progress, as it stands today, might simply be accelerated by removing the profit motive from industrial organization. The solution to restraining business sabotage is not to allow the full expansionary dynamics of technological change to run independently of human and ecological needs, but to assert greater democratic control over the direction of technical change, including restraining technological advancement where it threatens our collective wellbeing.

Data Appendix

As noted in Section 3.1, to make the statistical results more robust, I used two methods of measuring Dominant Semiconductor Capital. Method 1 calculated Dominant Semiconductor Capital as the top ten firms listed in the Compustat database by market capitalization, and Method 2 calculated Dominant Semiconductor Capital as the top ten *percent* of firms listed in the Compustat database by market capitalization. Thus, for Method 2, the number of firms included

in Dominant Semiconductor Capital rises and falls with the overall number of listed firms in the database.

The results of Method 2, shown below, do not differ greatly from those of the first, with the exception of the relationship between the rate of change in differential return on equity and rate of change in capital expenditures (see Figure 9). Under Method 2, there is no correlation between these two measures. Determining the specific reasons for these results are beyond the scope of the paper. However, the discrepancy highlights how subjective choices shape empirical results. Neither method of measurement represents the ‘true’ representation of Dominant Semiconductor Capital, but each serve as one possible quantitative proxy for much more complex social formations. While a more robust inquiry might add even more alternative measures and tests— such as grouping Dominant Semiconductor Capital by revenue—such measures must always be judged in light of qualitative evidence and with an awareness of the role of theory choice in shaping and limiting the interpretation of the data.

Results (Method 2)

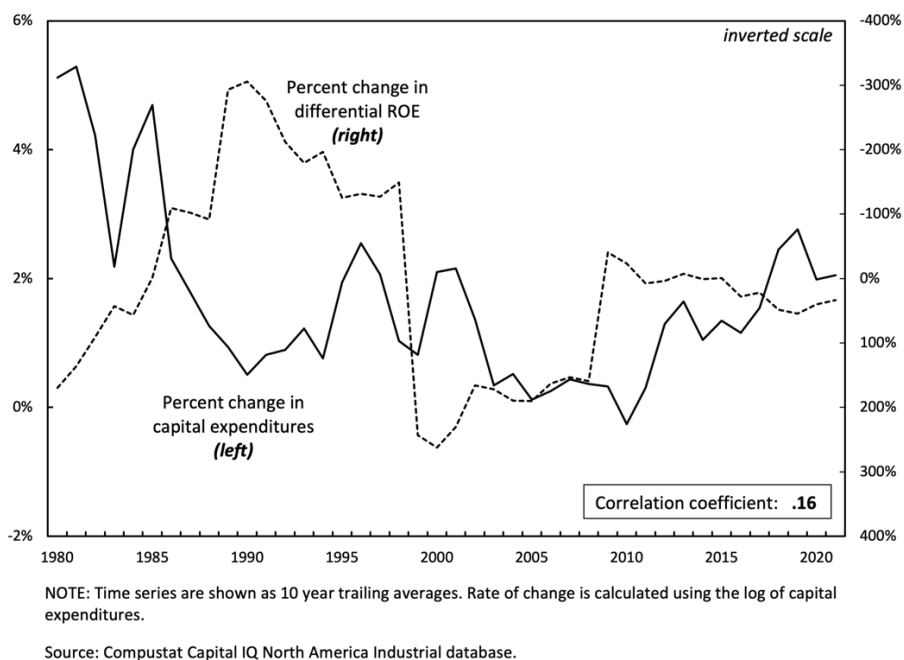


Figure 9. Percent change in differential return on equity (profit) and percent change in capital expenditures, Dominant Semiconductor Capital (Method 2)

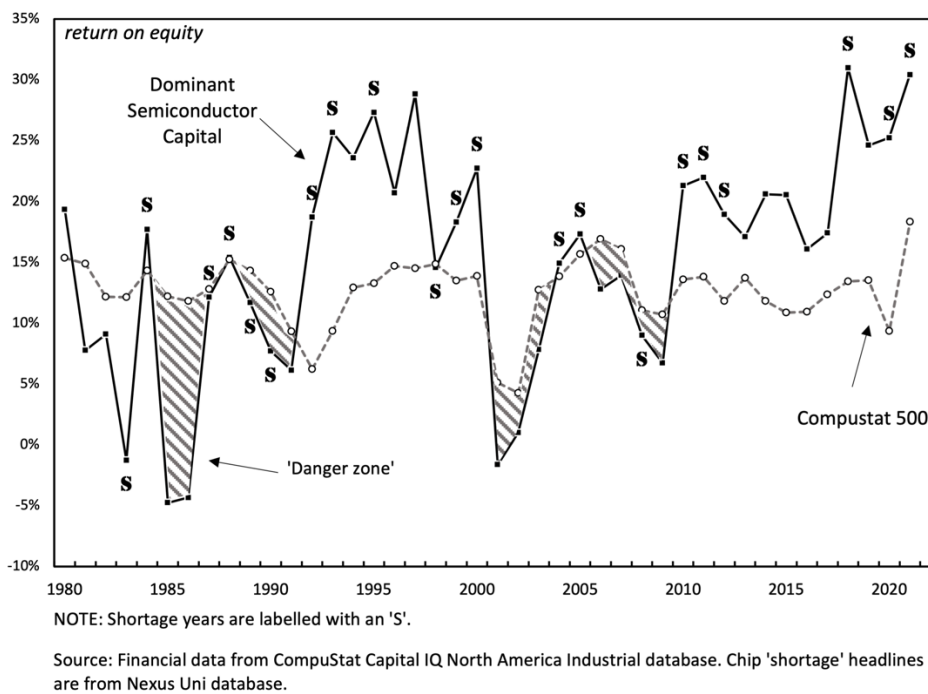


Figure 10. Return on equity and semiconductor shortages, Dominant Semiconductor Capital and Compustat 500 average (Method 2)

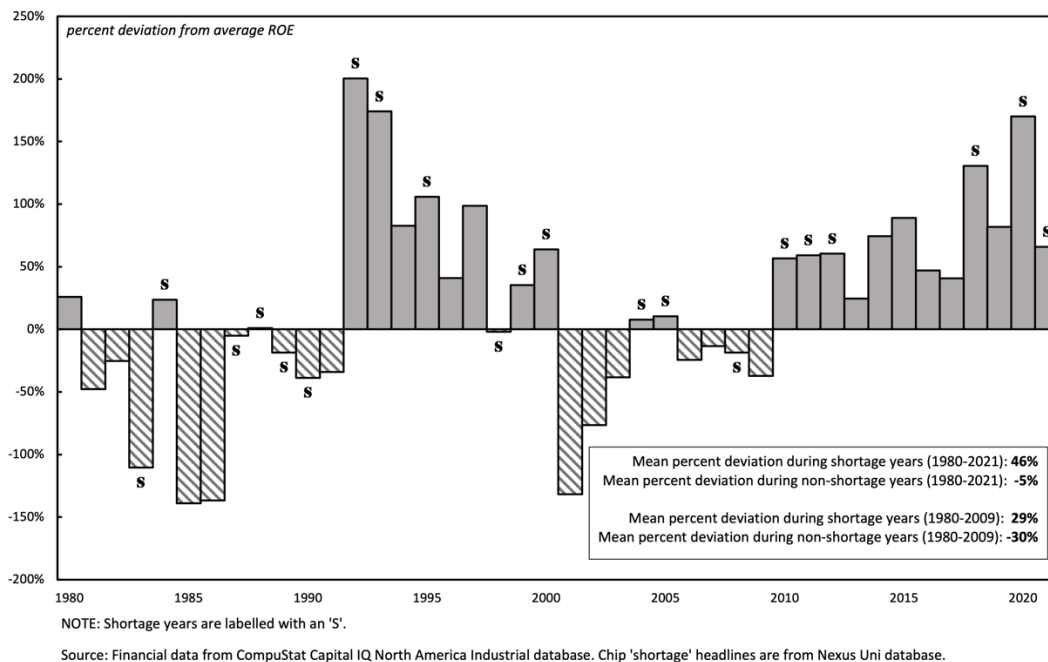


Figure 11. Differential profitability and semiconductor 'shortages' (Method 2)

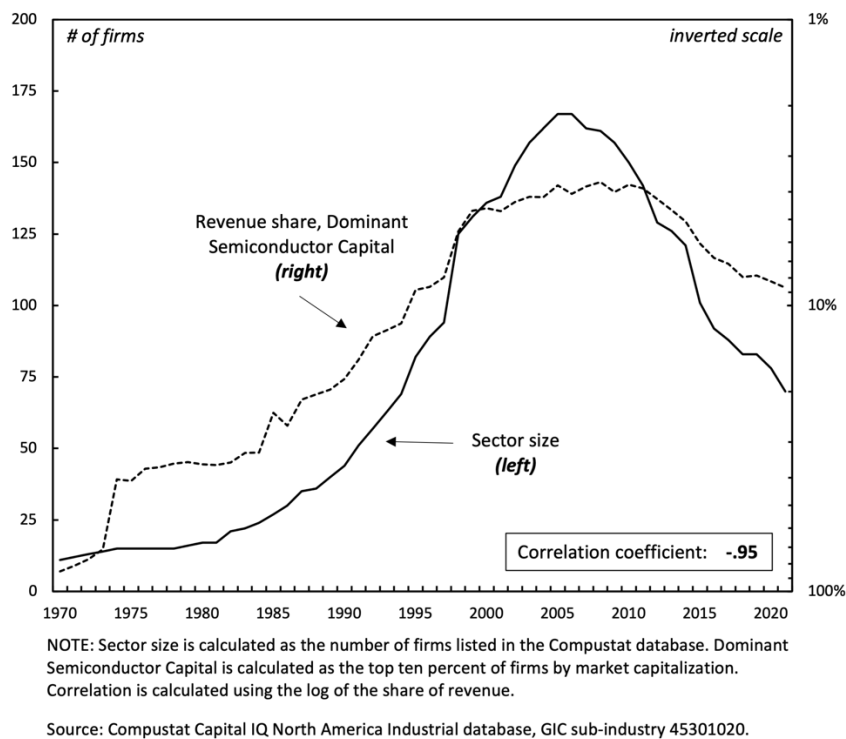
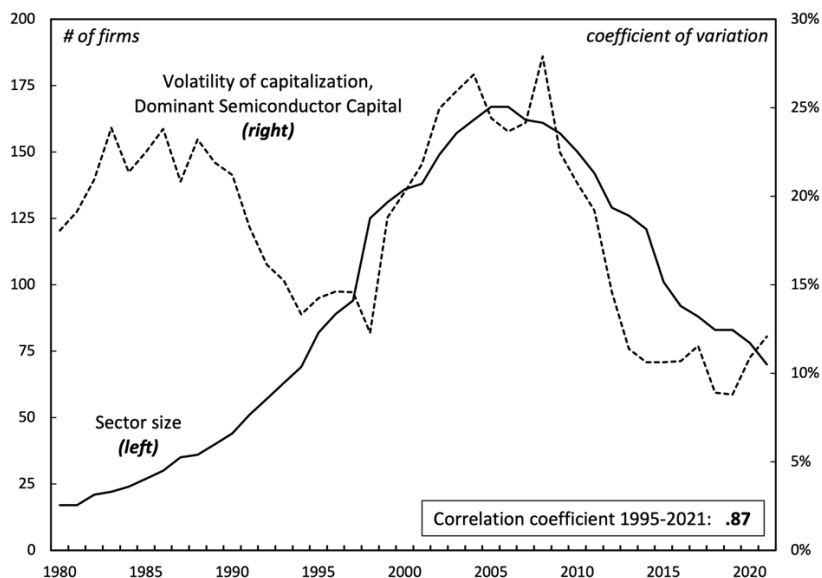


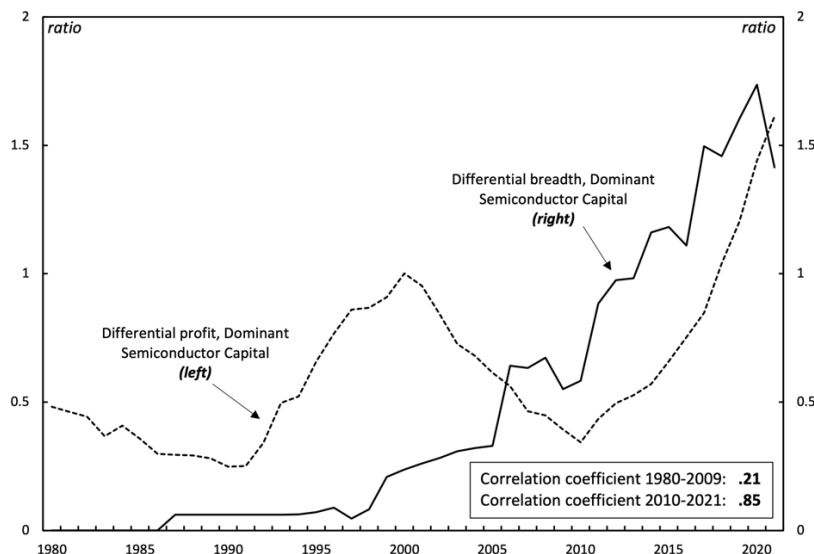
Figure 12. Sector size and Dominant Semiconductor Capital's share of the revenue (Method 2)



NOTE: The size of the sector is measured as the number of listed firms in the Compustat database. Dominant Semiconductor Capital is measured as the top ten percent of firms by market value. The volatility index is constructed by calculating the coefficient of variation (the standard deviation from the mean) of the rate of change in the average capitalization of the top ten percent of semiconductor firms. The volatility index is shown as a 10 year trailing average.

Source: Compustat North America Industrial Database.

Figure 13. Sector size and capitalization volatility, Dominant Semiconductor Capital (Method 2)



NOTE: Differential profit of Dominant Semiconductor Capital is calculated as the average annual net income of the top ten percent of semiconductor manufacturing firms by market capitalization, divided by the average net income for the Compustat 500. Differential breadth is calculated by dividing the ratio of annual acquisitions to annual capital expenditures for the top ten percent of firms by the same ratio for the Compustat 500. Series are shown as ten year trailing averages.

Source: Compustat Capital IQ North America Industrial database.

Figure 14. Differential buy-to-build and differential profitability, Dominant Semiconductor Capital (Method 2)

Bibliography

- Bichler, Shimshon and Jonathan Nitzan. 2013. "Francis' Buy-to-Build Estimates for Britain and the United States: A Comment." *Review of Capital as Power*, Vol. 1, No. 1, pp. 73-78.
- Bichler, Shimshon and Jonathan Nitzan. 2021. "The 1-2-3 Toolbox of Mainstream Economics: Promising Everything, Delivering Nothing." *Real-World Economics Review*. No. 98. December. pp. 23-48.
- Caligaris, Gastón, and Guido Starosta. 2019. "Revisiting the Marxist Skilled-Labour Debate." *Historical Materialism* 27, no.1: 55-91.
- Capri, Alex. 2020. "Semiconductors at the Heart of the US-China Tech War: How a New Era of Techno-Nationalism is Shaking up Semiconductor Value Chains." The Hinrich Foundation.
- Chu, Ming-chin Monique. 2013. *The East Asian Computer Chip War*. New York: Routledge.
- Cockshott, Paul, Shimshon Bichler, and Jonathan Nitzan. 2010. "Testing the Labour Theory of Value: An Exchange." The Bichler and Nitzan Archives. <http://bnarchives.yorku.ca/id/eprint/308>.
- Cockshott, W. Paul, and Allin Cottrell. 2005. "Robust correlations between prices and labour values: a comment." *Cambridge Journal of Economics* 29: 309-316.
- Duménil, G. 1983. "Beyond the Transformation Riddle: A Labour theory of Value." *Science and Society* 47, no.4: 427-250.
- Flamm, Kenneth. 1996. *Mismanaged Trade? Strategic Policy in the Semiconductor Industry*. Washington, DC: Brookings Institution.
- Flannery, Russell. 2022. "Global Chip Supply-Demand Imbalance Is Spurring Industry Change: NXP CEO Kurt Sievers." Forbes.com, May 25, 2022. <https://www.forbes.com/sites/russellflannery/2022/05/25/global-chip-supply-demand-imbalance-is-spurring-industry-change-nxp-ceo-kurt-sievers/?sh=24421ac222f5>.
- Foley, Duncan K. 1982. "The Value of Money, the Value of Labor Power, and the Marxian Transformation Problem." *Review of Radical Economics* 14, no.2: 37-47.
- Francis, Joseph A. 2013. "The Buy-to-Built Indicator: New Estimates for Britain and the United States." *Review of Capital as Power*, Vol. 1, No. 1, pp. 63-72.
- Grindley, Peter, and David Teece. 1997. "Managing Intellectual Capital: Licensing and Cross-Licensing in Semiconductors and Electronics." *California Management Review* 39, no. 2: 8-41.

- Harper, Michael. 1997. "Estimating Capital Inputs for Productivity Measurement: An Overview of Concepts and Methods." Paper prepared for the Conference on Measuring Capital Stock, Canberra, Australia, March 10-14, 1997. <https://www.oecd.org/sdd/na/2666894.pdf>.
- Hart, David M. 2001. "Antitrust and Technological Innovation in the US: Ideas, Institutions, Decisions, and Impacts, 1890–2000." *Research Policy* 30, no. 6: 923–936.
- Hunt, E.K., and Mark Lautzenheiser. 2011. *History of Economic Thought: A Critical Perspective*. 3rd ed. London: Routledge.
- Irwin, Douglas. 1996a. "The U.S.-Japan Semiconductor Trade Conflict." In *The Political Economy of Trade Protection*, edited by Anne Krueger, 5-14. Chicago: University of Chicago Press.
- Irwin, Douglas. 1996b. "Trade Policies and the Semiconductor Industry." In *The Political Economy of American Trade Policy*, edited by Anne Krueger, 11-72. Chicago: University of Chicago Press.
- Kelly, Leanna. 2022. "Global Semiconductor Chip Shortage Update." Investment U, March 14, 2022. <https://investmentu.com/chip-shortage-update/>.
- Malerba, Franco. 1985. *The Semiconductor Business*. London: Franco Pinter.
- Malone, Michael. 2014. *The Intel Trinity: How Robert Noyce, Gordon Moore, and Andy Grove Built the World's Most Important Company*. New York: Harper Business.
- Martin, Ulf. 2019. "The Autocatalytic Sprawl of Pseudorational Mastery." *Review of Capital as Power* 1, no.4: 1-30.
- McClellan, Bill. 2011. "Research Bulletin: Tracking the Top 10 Semiconductor Sales Leaders Over 26 Years." *IC Insights*.
- Morishimi, M., and F. Seton. 1961. "Aggregation in Leontief Matrices and the Labour Theory of Value." *Econometrica* 29, no.2: 203-220.
- Moseley, Fred. 2012. "A Critique of the Marginal Productivity Theory of the Price of Capital." *Real-World Economics Review*, no. 59: 131-137. <https://rwer.wordpress.com/2012/03/12/rwer-issue-59/>.
- Mumford, Lewis. 1971. *The Pentagon of Power*. London: Secker and Warburg.
- Nenni, Daniel, and Paul McLellan. 2013. *Fables: The Transformation of the Semiconductor Industry*. CreateSpace Independent Publishing Platform.
- Nitzan, Jonathan. 1992. "Inflation as Restructuring: A Theoretical and Empirical Account of the U.S. Experience." Unpublished PhD dissertation.

- Nitzan, Jonathan, and Shimshon Bichler. 2002. *The Global Political Economy of Israel*. London: Pluto.
- Nitzan, Jonathan, and Shimshon Bichler. 2009. *Capital as Power: A Study of Order and Creorder*. *RIPE Series in Global Political Economy*. New York and London: Routledge.
- Noble, David. 1979. *America by Design: Science, Technology, and the Rise of Corporate Capitalism*. Oxford University Press.
- Popper, Karl. [1963] 2002. *Conjectures and Refutations*. New York: Routledge.
- Robinson, Joan, and John Eatwell. 1973. *An Introduction to Modern Economics*. Maidenhead, UK: McGraw-Hill.
- Rowthorn, Bob. 1974. "Skilled Labour in the Marxist System." *Bulletin of the Conference of Socialist Economists* 8: 25-45.
- Shaikh, Anwar M., and E. Ahmet Tonak. 1994. *Measuring the Wealth of Nations. The Political Economy of National Accounts*. Cambridge: Cambridge University Press.
- Tilton, John. 1971. *The International Diffusion of Technology: The Case of Semiconductors*. Washington D.C.: The Brookings Institute.
- Veblen, Thorstein. 1908. "On the Nature of Capital: Investment, Intangible Assets, and the Pecuniary Magnate." *Quarterly Journal of Economics* XXII (1, November): 104-136.
- White, Edward, and Kathrin Hille. 2021. "Samsung warns of deepening semiconductor shortage." *Financial Times*, March 17, 2021. <https://www.ft.com/content/7e5e2d0f-3aa4-412b-898c-15acda4adb2e>.
- Wolff, Edward N. 1975. "The Rate of Surplus Value in Puerto Rico." *Journal of Political Economy* 83, no.5 (October): 935-950.