International Journal of Industrial Organization 69 (2020) 102565

Contents lists available at ScienceDirect

International Journal of Industrial Organization

journal homepage: www.elsevier.com/locate/ijio

The role of design method and process technology in stable outsourcing equilibria☆

Jin-Hyuk Kim[a,*], Takehiko Komatsu[b], Hideo Owan[c]

[a] Department of Economics, University of Colorado at Boulder, 256 UCB, Boulder, CO 80309, USA
[b] Amazon Japan G.K., Arco Tower Annex 1-8-1 Shimomeguro, Meguro-ku, Tokyo 153-0064 Japan
[c] Faculty of Political Science and Economics, Waseda University, 1-6-1 Nishiwaseda, Shinjuku-ku, Tokyo 169–8050, Japan

ARTICLE INFO

Article history:
Received 1 March 2019
Revised 9 December 2019
Accepted 17 December 2019
Available online 24 December 2019

JEL classification:
L23
L63
M20

Keywords:
Outsourcing
Vertical integration
Semiconductor industry

ABSTRACT

We analyze a model of vertical (dis)integration between manufacturing and design in a monopolistically competitive market. Specialized input manufacturers can serve multiple design firms and the manufacturer-designer pairs negotiate a non-binding contract to share input customization cost and production surplus. Hand-collected data on 387 product lines from 118 semiconductor firms are used to predict the firm's decision to outsource manufacturing. We find that, for instance, the use of design tools that facilitate collaboration and process technologies that facilitate learning are both positively associated with outsourcing, consistent with the model's prediction.

1. Introduction

With the technological advancement in the 20th century and the establishment of such pioneering firms as Fairchild Semiconductor, practicable semiconductor devices, especially in the form of integrated circuits (ICs, or 'chips'), have shaped the evolution of electronics industries. Originally, semiconductor devices were produced by the so-called integrated device manufacturers (IDMs) having both the capacity to design and manufacture semiconductor devices. A well-known example is Intel, which developed the first microprocessor in 1971 and still dominates the microprocessor product market today.

By the late 1980s, however, the so-called foundry model came into being. The foundry model refers to the separation of chip design and fabrication process into different business entities.[1] Design firms are often called 'fabless' and specialized manufacturers are called 'foundries.' Concurrently, most of the foundries were established in Asia, such as Taiwan Semiconductor Manufacturing Company (TSMC, founded in 1987), and there are numerous fabless companies today in many parts of the world specializing in chip designs for a broad range of applications from mobile phones to internet-of-things to self-driving cars.

☆ We would like to thank the Editor, Stephen Ryan, and two anonymous referees for helpful comments and suggestions.
* Corresponding author.
E-mail addresses: jinhyuk.kim@colorado.edu (J.-H. Kim), takkoma@amazon.co.jp (T. Komatsu), owan@waseda.jp (H. Owan).

[1] Integrated device manufacturers have spun off their foundry divisions, so that they do not compete against their foundry customers in the same product market. On the other hand, a 'pure-play' foundry does not sell semiconductor devices of its own design. For the purpose of this paper, we do not distinguish these two types of foundries.

https://doi.org/10.1016/j.ijindorg.2019.102565
0167-7187/

Although a number of industry reports (such as McKinsey & Co's) and newspaper articles have talked about what might have given rise to the popularity of the foundry model, there have been relatively few studies that rigorously investigate such arguments theoretically and/or empirically, especially using data from the 21st century (see below). This paper fills this gap by presenting an industry equilibrium that explains the conditions under which a stable outsourcing equilibrium may arise and empirically showing that the factors that are highlighted in the model as the drivers can indeed explain a significant variation in the data.[2]

The central argument behind the emergence of the foundry model among industry experts is that the development of design tools and standardization of process technologies played a key role in facilitating the specialization in the vertical chain (e.g., Fuller et al., 2003; Macher and Mowery, 2004; Saito, 2009). Further, as the applications for semiconductor devices became more sophisticated over time, the capital requirements for developing new chip designs and process technologies became too large for some IDMs. Thus, some firms decided to focus on chip design while others on fabrication given their relative capabilities and strength.

To be precise, the chip design practice was facilitated by the development of electronic design automation (EDA) tools and cell libraries, which helped the layout and interface construction as well as performance simulation of novel chip designs. At the same time, the so-called "complementary metal oxide semiconductor" (CMOS) processes emerged as a prominent manufacturing process. The CMOS process is standardized, hence, applicable to a variety of types of semiconductor products, so the foundries could aggregate demands from multiple design firms, process on bulk standard CMOS wafers, and achieve high yield improvements.[3]

Since yields determine productivity and the CMOS process is more versatile (i.e., can be used to manufacture a wide range of products) than others (e.g., the production process for analog logic devices), the CMOS process tends to benefit from both economies of scope and scale. For instance, TSMC succeeded greatly by developing and launching CMOS processes for smart-phones as well as other customers in the communications industry. Other foundries follow more or less the same pattern of launching a new process, pooling demands and ramping up production until they launch another process and further expand their customer base.[4]

We aim to contribute to the literature on vertical (dis)integration by incorporating the above arguments into an economic model of outsourcing and presenting some systematic evidence on it. In this regard, this paper is consistent with and complementary to the management literature on the semiconductor industry based on detailed qualitative studies. For instance, Linden and Somaya (2003) put forward arguments and facts that are broadly consistent with ours in that the CMOS process and the EDA tools helped develop the vertical disintegration of the supply chain and the modular production of various products, especially system-on-a-chip.

The literature on outsourcing that is closest to our approach is the transaction cost economics (TCE) literature (e.g., Coase, 1937; Williamson, 1975), on which most of existing work on the "make-or-buy" decision is based.[5] The dominant prediction of the TCE literature is that specific assets create an ex-post opportunism (or hold-up of quasi-rents), so firms tend to internally produce outputs rather than outsource (Williamson, 1985). We build on this framework a monopolistic competition model and empirically measure the asset or technology specificity at a finer level of product lines than existing papers do.

For instance, Monteverde (1995) is a seminal paper on the semiconductor industry, where technical dialog (i.e., unstructured design-fabrication interaction) is used to explain the "curious phenomenon of fablessness." It was argued that analog and memory chips require close coordination between design and process engineers while logic chips are less likely to do so because of automated design tools and documented process rules. To confirm this claim, industry experts were asked to rate each of the sample *firms* on the relative magnitude of technical dialog necessary for the product strategy each company was pursuing.[6] It was found that, when the volume of technical dialog is large, integration was more likely than outsourcing.

Leiblein and Miller (2003) embed the technical dialog put forward by Monteverde (1995) into a similar empirical framework. The novelty is that they expand the dataset both in terms of sample size and the unit of analysis.[7] Specifically, their sample comes from a 1996 survey of global IC manufacturers (with 117 responding firms), where the unit of observation is defined by a combination of seven product categories (analog; application-specific integrated circuit (ASIC); discrete; digital

[2] Our model is not specific to semiconductor industry but it may be applicable to a wider set of industries. Hence, our analysis complements those of other scholars who have studied outsourcing/integration decisions in such industries as aerospace, automobile, and chemical industries (e.g., Monteverde and Teece, 1982; Masten, 1984; Lieberman, 1991).

[3] Gruber (1992) showed that the learning curve in the production of memory chips is associated with certain types of memory products that use CMOS process. Specifically, the learning curve was steep for Programmable Read Only Memory which was well integrated into CMOS process, but not for Dynamic Random Access Memory which was harder to integrate.

[4] Foundries do not necessarily focus on leading-edge processes as the traditional IDMs do, because there are typically tradeoffs between high speed and power consumption, and they put more emphasis on standardized and modular processes in order to satisfy the changing needs of the fabless customers. Thus, foundries use CMOS or CMOS-based modified processes.

[5] We do not attempt to survey the literature on TCE here (see, e.g., Tadelis and Williamson (2012) for a survey on TCE). Below, we only discuss those papers that use data from the semiconductor industry and focus on the technological aspects as we do here. We also acknowledge that there are other views based on historial capacities and internal resources.

[6] Monteverde (1995)'s sample comprises 23 public semiconductor firms whose stocks were trading on a major U.S. exchange in 1991. Monteverde controls for other factors such as patent counts and firm size proxies (although these were not significant), which we do similarly below.

[7] Another difference is that Leiblein and Miller (2003) include the resource-based view as well as real option value by including production capacity and the number of sourcing relationships and product subfields. We partly address the resource-based view using patent counts.

signal; memory; microprocessor; telecommunication) and manufacturing process nodes (from 1- to 0.25-micron). Similar to Monteverde (1995), asset specificity is coded high for analog, ASIC, and memory and low for all other categories.

We follow this stream of literature by refining the data sample on a couple of dimensions. First, our sample comes directly from a trade association, covering 118 semiconductor firms from around the world, and we hand-collected reported data on technical specifications of the firms' products from data sheets and company brochures. Second, we classify product lines into five main categories (analog; discrete; logic, memory; micro) and 19 subcategories. The latter gives us a richer variation to exploit because each product line can have a different set of technical specifications even within the same main category, and firms can also make a different sourcing decision across product lines.

Since both Monteverde (1995) and Leiblein and Miller (2003) code asset specificity at a higher level category, it means that the contribution of underlying design and process technologies are not clearly identified if there are nontrivial variations within main product categories. Thus, to refine the contribution of these technological factors, we would separately code the use of EDA tools and the adoption of CMOS processes for all our sample product lines, which vary across product subcategories. By including these variables in addition to the (traditional) product category dummies, we are better able to identify the effects of the said factors.

Macher (2006) provides the closest analysis to ours in this regard by using data on 179 process groups from 36 manufacturing facilities that started between 1995 and 2001. While the focus is on how organizational structure (i.e., specialized versus integrated manufacturers) moderates the effects of TCE (which is proxied at the category level; e.g., analog and memory chips have ill-structured problems while logic is not) on process performance (such as development time and die yields), one of their results shows that the EDA tools and the market share held by CMOS technologies at the start of each process affect the outsourcing decision.

To our knowledge, the literature has not integrated the empirical evidence with a model of industry equilibrium that illustrates the effects of input customization and learning curve. Further, the foundry-fabless model has become increasingly important in the global semiconductor industry in the mid-2000s when the Moore's Law was slowing down (e.g., Waldrop, 2016).[8] More firms have since diversified their business opportunities rather than necessarily moving to leading-edge processes. Thus, our paper contributes to the literature by bringing an outsourcing model to recent 2007 data when the foundry model became more prominent.[9]

The rest of the paper is organized as follows. Section 2 presents the model of outsourcing and shows the stability of market equilibria. Section 3 describes the dataset and documents the empirical evidence. Section 4 concludes.

2. Theoretical analysis

2.1. Model

To understand how the design methods and process technologies can explain the transition from IDM-dominated industry structure to the foundry model, we first analyze a theoretical model that embeds a stable bargaining outcome into the industry outsourcing model proposed by Grossman and Helpman (2002, 2005). The difference is that we allow the input supplier and the final producer to share the customization cost, as was emphasized by industry executives.[10] We then look for the bargaining solution between multiple final producers and the common manufacturer (input supplier) that is stable against ex-post renegotiation.

To be more precise, Grossman and Helpman (2005) propose a model of outsourcing that involves a search for a partner where relationship-specific investments are governed by incomplete contracts. In their model, final producers cannot manufacture input, so they must outsource this activity to specialized input suppliers. While our focus here is on the integration versus outsourcing decision, our model shares the feature that the final producer is matched with the closest input supplier and also the supplier can serve multiple final producers. We refine their Nash Bargaining solution with a stable bargaining solution, but we do not incorporate a fixed search cost to find a partner as they do.

In another related paper, Grossman and Helpman (2002) analyze the integration versus outsourcing decision, but the main mechanisms are different. In their model, the matching between input suppliers and final producers are realized by a random matching function, and the input supplier suffers from the buyer's holdup problem, which gives a reason for vertical integration. In their baseline model, there is no input customization cost. In an extension, they allow the input supplier to optimally choose the degree of input specialization on the interior of a unit circle, but unlike in our model, the final producer incurs the entire cost of modifying the supplier's input for their use, instead of sharing the cost with the supplier.

[8] Gorden Moore, a co-founder of Intel, said that the number of transistors that can be fitted into a single chip would double roughly every two years, resulting in faster performance and lower cost. This development schedule has been held true until around 2004.

[9] The semiconductor industry has been studied by a number of researchers; however, most of the studies have focused on Moore's law and the associated productivity growth (e.g., Jovanovic and Rousseau, 2002; Flamm, 2003; Aizcorbe and Kortum, 2005).

[10] For instance, executives we interviewed commented: "We invest a large amount of high quality engineering team to understand how the [process] technology actually works", "Instead of simply providing specifications, we look for small modifications to their manufacturing capability", and "We build those limitations into our design methodology."

We consider a final goods market in which firms produce differentiated products and choose the mode of input production. On the demand side, there is a representative consumer who maximizes a constant elasticity of substitution (CES) utility function $\left(\sum_{i=1}^{n} y_i^{\alpha}\right)^{1/\alpha}$, where y_i is the consumption of variety $i = 1, \ldots, n$ and $\alpha \in (0, 1)$ measures the degree of substitutability between the final goods (Dixit and Stiglitz, 1977). As is well-known, the CES preference yields the demand for each variety as follows:

$$y_i = A p_i^{-1/(1-\alpha)}, \tag{1}$$

$$A = \frac{E}{\sum_{i=1}^{n} p_i^{-\alpha/(1-\alpha)}}, \tag{2}$$

where p_i is the price of good i and E is the aggregate expenditure or the market size, which we hold constant. Final producers take A as given.[11]

On the supply side, each final good is associated with a point on the circumference of a unit circle, so that the location of the good represents the specification of the input required to produce variety i. Without much loss, we assume a one-to-one mapping between product location and input specification, so that an input is useless for the product unless it exactly matches the specification. Further, one unit of input is required to produce one unit of each product, and besides the intermediate good there is no other input in the final good production.

Production of variety i requires a fixed investment f_d in product design. Final goods can be produced by vertically integrated firms or by specialized design firms that purchase customized inputs from the manufacturers. Input suppliers make a fixed investment f_m in the manufacturing process, which is optimized to a point on the unit circle. Given the trend of increasing facility costs, we assume that the cost of setting up manufacturing facilities is larger than the cost of designing a product, so that each supplier may serve multiple design firms.

The design firms and input suppliers randomly locate themselves around the circle, and they observe one another's location perfectly. Thus, when a design firm outsources the production of the intermediate good, it will approach the closest input supplier with a distance $x \in (0, 1/2]$ along the unit circle. In that case, the designer and the supplier must coordinate to customize the input (e.g., develop a prototype).[12] We assume that the customization process is unique for each design, and the cost of customization is proportional to the distance, that is, ηx.

The designer-cum-final-producer and the manufacturer-cum-input-supplier share the cost of customization. Let λ be the share of the cost the design firm bears while the manufacturer's share is $1 - \lambda$. We assume that the cost incurred during the input customization is observable to both parties but it is subject to the hold-up problem. Hence, the parties must bargain over the cost sharing given the continuation value when the prototype is developed. If they cannot reach an agreement at this stage, then we assume that the design firm has to exit the market.[13]

Once a prototype is developed, both parties have a double coincidence of wants concerning the production of the final good. That is, we assume that the parties will choose the efficient quantity y_i to maximize the total surplus and bargain over the division of this surplus using a lump-sum transfer T_i. If the bargaining breaks down at this stage, then the designer firm must exit the market (where we assume a zero outside option value). To summarize, in the case of outsourcing, bargaining over λ and T takes place before and after input customization, respectively.

Given the holdup problem in our model and the multiple design firms per supplier, we adopt the stable bargaining outcome of Stole and Zwiebel (1996) as our solution concept.[14] A stable bargaining outcome deals effectively with our bargaining problem and posits that the parties in any pairwise negotiation split the difference equally relative to their outside options and no party can improve its payoffs through renegotiation. Once a stable bargaining solution is reached, the parties would make required investments and transfer payments as agreed upon.

On the other hand, when a vertically integrated firm enters the market, it pays the sum of the two fixed costs, $f_d + f_m$. However, since it makes its own input, the location of input design and manufacturing expertise is perfectly aligned, so there is no additional cost for developing a prototype.[15] Importantly, we assume that an integrated firm does not work with

[11] The rationale is that the number of products, n, is sufficiently large, so the second-order effect through A is ignored. We think that this is a reasonable model to describe the semiconductor market given the variety of products. Allowing for the second-order effect would make the solution rather intractable; we conjecture that it would not affect our qualitative results.

[12] For instance, EDA produces a standardized design that relies on the use of libraries of components for their production. Such designs are generally not optimized for specific foundries, especially as firms need to use the libraries that contain increasingly complex functionalities, so they require some joint development work prior to production.

[13] A design firm may invest in manufacturing capacity and become an integrated firm if it fails to reach an agreement with the closest supplier. We abstract from this possibility because it complicates our exposition of the bargaining solution without real changes in qualitative results.

[14] Stole and Zwiebel (1996) deal with the bargaining between an employer and multiple employees in the context of non-binding contracts, and propose a solution for which there exists an extensive-form game whose unique subgame-perfect equilibrium has the same structure as the Shapley value.

[15] One can think of this as a reduced form of an internal optimization process within integrated firms, whereby there can be an additional cost of customizing manufacturing process to match the required design specification. As long as this cost is far smaller than that for the designer-manufacturer pair, none of the following analysis and qualitative results would be affected.

a specialized designer and vice versa. This is because the prototype of the designer would be revealed to the integrated firm, and there is a risk of misappropriation unless the production unit is spun off.[16]

There is considerable evidence that learning curve matters in the semiconductor industry (e.g., Irwin and Klenow, 1994; Gruber, 1998). One advantage of outsourcing relative to vertical integration is that the manufacturer can aggregate demands from multiple design customers, which tends to lower the marginal cost of production (e.g., Riordan and Williamson, 1985). In particular, when different products use similar process technologies, manufacturers can not only improve the production yield but also re-use the additions and options made for other products.[17]

We thus assume that for specialized input suppliers there exist economies of scope/scale from collaborating with multiple design firms.[18] For instance, foundries can combine fast- and flexible-ramp products to reduce down time in production, and different products can be used to identify problems in the production process faster. Specifically, let c_I and $c_S(m)$ denote the marginal cost of an integrated firm and a specialized manufacturer, respectively, where $m \geq 1$ is the number of customers per manufacturer. We then assume that the manufacturer's marginal cost, $c_S(m)$, satisfies $c_S(1) = c_I$, $dc_S(m)/dm < 0$, $d^2c_S(m)/dm^2 \geq 0$.

Finally, the timing of the game is as follows. First, firms simultaneously enter as a vertically integrated firm, a specialized designer, or an input manufacturer. Second, design firms bargain with the closest supplier to customize the input. Third, producer and supplier bargain over the quantity of production and a lump-sum payment. Fourth, designers and integrated firms set prices and consumption occurs. We focus on the symmetric equilibrium where firms on the same side (i.e., final producers and input suppliers) are located at equidistant intervals on the unit circle.

2.2. Stable equilibria

We will first show that there are a multiplicity of equilibria in our model, and then illustrate the stable industry equilibrium in a tâtonnement process, which gives rise to a two-dimensional phase diagram as in Grossman and Helpman (2002). That is, firms enter the market when they expect to earn positive profits and exit when they expect losses. However, one difference is that our definition of "stable" equilibria incorporates the stable post-entry bargaining outcome à la Stole and Zwiebel (1996), which is necessary given the economies of scope/scale.

As usual, we solve for the equilibrium using backward induction. The consumer's preference implies a fixed mark-up pricing ($p_i^*(m) = c_S(m)/\alpha$) for the design firms, where m is a designer-to-manufacturer ratio. From (1), the market demand for each variety i is $y_i^*(m) = A(c_S(m)/\alpha)^{-1/(1-\alpha)}$. Similarly, integrated firms set $p_i^* = c_I/\alpha$ for their final goods; and the resultant demand is $y_i^* = A(c_I/\alpha)^{-1/(1-\alpha)}$. We assumed that design firms and input suppliers reach an efficient agreement, so they would choose the profit-maximizing quantity $y_i^*(m)$ and share the profits.

From (2), it follows that the level of market demand is

$$A = \frac{E}{n^D(\alpha/c_S(m))^{\alpha/(1-\alpha)} + n^I(\alpha/c_I)^{\alpha/(1-\alpha)}}, \tag{3}$$

where n^D and n^I are the number of design and integrated firms, respectively.

Let $S(m)$ denote the share of the profit captured by the manufacturer who produces intermediate goods for m (symmetric) designers. Similarly, let $B_i(m)$ denote the share of the profit captured by the designer of product i when its partner manufacturer produces customized inputs for m different designers. Then,

$$\begin{aligned} S(m) &= m[T_i - c_S(m)y_i^*(m)], \\ B_i(m) &= p_i^*(m)y_i^*(m) - T_i. \end{aligned}$$

Let $V(m)$ denote the total surplus created by the collaboration, that is, $V(m) = S(m) + mB_i(m) = m[p_i^*(m) - c_S(m)]y_i^*(m)$. A stable bargaining outcome requires that for any pairwise relationship between a manufacturer and a designer, the net surplus relative to the party's outside option is split equally between the two parties. Therefore, the bargaining outcome must satisfy $S(m) - S(m-1) = B_i(m)$, the solution of which can be characterized as follows.

Lemma 1. *The stable bargaining outcome over the division of the surplus (after a prototype is developed) is that each design firm pays its input supplier* $T_i^*(m) = c_S(m)y_i^*(m) + \sum_{k=1}^{m} \frac{V(k)}{m(m+1)}$.

If there were no learning effect ($dc_S/dm = 0$), then each bilateral relationship yields the same amount of surplus in the symmetric equilibrium, so the stable bargaining outcome would lead to an equal sharing of the surplus, $V(1)/2$. However,

[16] Some IDMs (e.g., Samsung, Advanced Micro Devices) spun off an independent foundry business; otherwise, design firms would be unwilling to share sensitive design information. For instance, a number of online articles mention that Intel, an IDM, never really had a foundry business because fabless customers face the risk of IDMs taking their designs or not prioritizing their chips.

[17] For instance, interviewees said "If you double the amount of material you process, your yield improves twice faster" and "Particularly with both using similar kinds of technology with small modification, there will be a shared development and time-shifted demand on the factory. It improves the profitability of the facility quite considerably."

[18] For our modeling purposes, we cannot clearly distinguish the economies of scope and economies of scale. According to our interviewees, however, economies of scale may take place first with a relatively few initial customers, followed by economies of scope when the foundry process becomes more mature and versatile.

with economies of scope/scale, there are inframarginal gains from taking on additional design firms; and the parties renegotiate the division of surplus given the hold-up problem, after a design firm joins or leaves the same manufacturer.[19]

Prior to developing a prototype, the design firm and its closest input supplier must negotiate the sharing of the input customization cost. Specifically, they enter into a non-binding agreement to make an investment of $\lambda\eta x$ and $(1-\lambda)\eta x$, respectively, where $\lambda \in [0, 1]$ denotes the cost share of the design firm. Anticipating the stable bargaining outcome once the investment in the prototype is sunk, the bargaining solution requires that the parties set λ to split the surplus equally; that is,

$$S^*(m) - S^*(m-1) - (1-\lambda)\eta x = B_i^*(m) - \lambda\eta x.$$

Thus, the stable bargaining outcome implies an equal split of the input customization cost, namely, $\lambda^* = 1/2$, and the collaboration between a design firm and its closest supplier is feasible if and only if

$$S^*(m) - S^*(m-1) + B_i^*(m) \geq \eta x.$$

Letting $\Delta(m) = S^*(m) - S^*(m-1) + B_i^*(m)$, a design firm that finds the closest supplier at a distance greater than $\Delta(m)/\eta$ cannot expect the negotiation on input customization to be successful. Thus, given that suppliers are equally distanced on a unit circle in the symmetric equilibrium, a designer-manufacturer pair will collaborate if the distance between them is less than $l(m) = \min\{\Delta(m)/\eta, 1/2n^M\}$, where n^M denotes the number of specialized input manufacturers.

When making an entry decision, a designer regards all equidistanced configurations of manufacturer locations as equally likely. This means that the distance between a designer and its closest manufacturer is a random draw from a uniform distribution with support $[0, 1/2n^M]$. Therefore, a designer expects to find a feasible supply relationship with a probability $2n^M l(m)$. Similarly, a manufacturer that enters can expect to serve a measure $2n^D l(m)$ of designers, because in any one direction the manufacturer can work with design firms located within a distance of $l(m)$.

In what follows, we restrict attention to the parameterization $\Delta(m)/\eta > 1/2n^M$, so that design firms can always find a supplier willing to share investment costs. Specifically, for any values of model parameters other than η, there exists a value $\bar{\eta} > 0$ such that the above inequality holds when $\eta < \bar{\eta}$.[20] If this condition is not satisfied, then design firms will be only able to find a collaborating supplier with a probability of $(2n^M\Delta(m))/\eta < 1$, and will be forced to exit without recovering the sunk entry cost with complementary probability.

However, our qualitative results would not change even if the above inequality does not hold. If η is sufficiently large, then an entrant will only be able to enter as an integrated firm. If the above inequality is not satisfied but η is relatively small, then similar results continue to hold because expected profits change only continuously while the zero profit conditions still need to hold. We thus believe that the above restriction on parameterization is innocuous and it helps us simplify our exposition that follows.

Equilibrium requires zero expected profits for each type of firms that enter in positive numbers. It follows that designers and manufacturers enter the market concurrently because if only one type of firms enters, then there is no profit to be made. Thus, the number of designers served by each manufacturer as well as the level of market demand in equilibrium is determined by the following zero expected profit conditions:[21]

$$S^*(m) - m(1-\lambda^*)\eta \mathsf{E}(x_i) - f_m = 0, \tag{4}$$

$$B_i^*(m) - \lambda^*\eta \mathsf{E}(x_i) - f_d = 0, \tag{5}$$

where $\mathsf{E}(x_i) = 1/4n^M$ is the expected distance between a designer and its closest supplier. Let (A_O^*, m^*, n^{M*}) denote the solution to (4) and (5); note that $n^{D*} = m^* n^{M*}$.

On the other hand, if an integrated firm enters, then the zero profit condition is

$$S(1) + B(1) - f_m - f_d = 0, \tag{6}$$

where there is no customization cost. Let A_I^* denote the solution to (6).

Proposition 1. *(a) A pure integration equilibrium ($n^{I*} > 0, n^{D*} = 0$) always exists and is globally stable. (b) A continuum of pure outsourcing equilibria ($n^{I*} = 0, n^{D*} > 0$) exist and are globally stable if and only if $A_O^* < A_I^*$. (c) A mixed equilibrium ($n^{I*} > 0, n^{D*} > 0$) exists and is saddle point stable if and only if $A_O^* < A_I^*$.*

[19] Firms are fully aware of the shared nature of the production process and that their project may have a beneficial effect on other clients of the same manufacturer. When asked about the foundry trying to meet the requirements of other customers, one executive said "There always is a conflict of interests. It's a contract and it's a negotiation [...] It is probably not a zero sum game."

[20] To see this, let $F(\alpha, c_I, c_S(\cdot), f_d, f_m, \lambda, \eta) = \min 2n^M\Delta(m)$, which is strictly positive because $n^M \geq 1$ and $\Delta(m) > 0$ for outsourcing equilibria to exist. Then, it can be shown that F is weakly decreasing in η because the convex closure of $n^M\Delta(m)$ is weakly decreasing in η in the strong set order. Thus, there exists a value $\bar{\eta} > 0$ such that $F > \eta$ when $\eta < \bar{\eta}$.

[21] Firms can therefore earn negative or positive profits ex post. There may be other sources of heterogeneity (such as productivity) that generate further ex-post profit differentials. However, we think that they are not necessarily central to the mechanism of our model, so we prefer to consider only the uncertainty regarding the location of the firms.

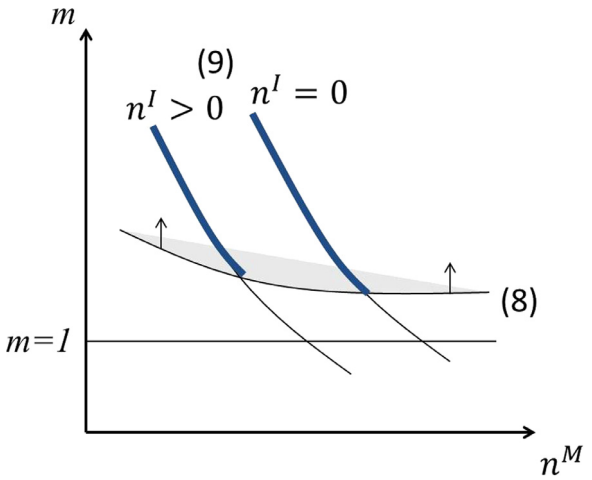

Fig. 1. Pure and mixed outsourcing equilibruia.

To illustrate the equilibrium with outsourcing, combine (4) and (5). Then the solution (A_O^*, m^*, n^{M*}) must satisfy

$$A_O^* c_S (m^*)^{-\frac{\alpha}{1-\alpha}} \left(\frac{1}{\alpha} - 1\right) \alpha^{\frac{1}{1-\alpha}} = \frac{\eta}{4n^{M*}} + \frac{f_m}{m^*} + f_d. \tag{7}$$

As shown in the proof of Proposition 1, a necessary and sufficient condition for outsourcing equilibria to be stable is $A_O^* < A_I^*$. Substituting A_I^* from (6) analogously, $A_O^* < A_I^*$ can be written as

$$\left(\frac{c_S(m^*)}{c_I}\right)^{\alpha/(1-\alpha)} < \frac{f_m + f_d}{\frac{\eta}{4n^{M*}} + \frac{f_m}{m^*} + f_d}. \tag{8}$$

On the other hand, combining (3) and (7) yields

$$E(1-\alpha) = \frac{\eta}{4} m^* + f_m n^{M*} + f_d n^{M*} m^*, \tag{9}$$

when $n^{I*} = 0$ (i.e., no integrated firm enters).

Notice that (8) is not satisfied at $m^* = 1$. Thus, if a given set of parameter values were to satisfy (8), then $m^* > 1$. Further, when (8) is satisfied, there is a continuum of (m^*, n^{M*}) pairs that may satisfy (9). This is illustrated in Fig. 1. Given the multiplicity of equilibria, it would be thus difficult to pinpoint the relative number of design firms versus input suppliers in an outsourcing equilibrium; however, (8) tells us under which conditions stable outsourcing equilibria exist.

We also note that if (8) is satisfied, then the representative consumer's indirect utility would be higher in an outsourcing equilibrium than in an integration equilibrium. This is because the indirect utility is inversely proportional to the level of market demand, A, holding constant other factors of the model. Since an outsourcing equilibrium is only stable when $A_O^* < A_I^*$, the outsourcing environment in a product market can be welfare improving in our model, absent other considerations and as long as the representative consumer model is valid.

2.3. Model prediction

To illustrate the emergence of stable outsourcing equilibrium, Fig. 2 depicts for a given supplier-to-manufacturer ratio m^* the phase diagram for the number of design firms (n^D) on the x-axis and the number of integrated firms (n^I) on the y-axis. Specifically, substituting in for A from (3), the OO-curve traces the combinations of (n^D, n^I) that are consistent with the level of market demand A_O^*, which is a function of n^M; and the VV-line does the same for the constant A_I^*. Hence, these are the combinations of (n^D, n^I) that result precisely in zero expected profits given A_O^* and A_I^* from above, respectively.

Since in an equilibrium design firms and input suppliers coordinate to enter the market with a ratio of m^*, the firms' belief on the marginal costs, $c_S(m^*)$ and c_I, are held fixed. However, A_O^* decreases as n^{M*} increases even if m^* is held constant. This is because initially a larger number of input suppliers lowers the input customization cost fast, which leaves a large enough demand level for more integrated firms to enter; however, as the number of suppliers keeps increasing, the benefit has a lower bound, so eventually n^I has to fall, as illustrated by the OO-curve.

By the same logic, the OO-curve is less steep than the VV-line. This implies that there is a mixed sourcing equilibrium (i.e., $n^D > 0$, $n^I > 0$) whenever pure outsourcing equilibria exist. However, the mixed equilibrium is saddle point stable, which implies that without some coordinating mechanism or an industrial policy to hop on the saddle path, the mixed equilibrium would not be stable. Thus, our main theoretical prediction concerning stable outsourcing equilibria is based on (8), where simple comparative statics predictions are as follows.

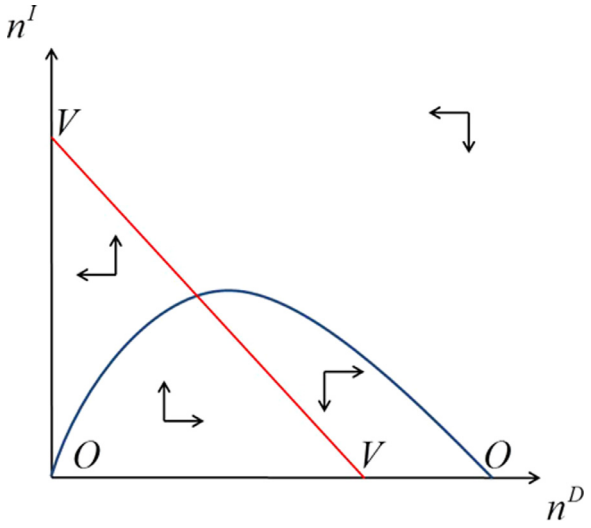

Fig. 2. Phase diagram of number of firms.

Corollary 1. *Holding everything else constant, stable outsourcing equilibria are more likely to exist if (i) η is smaller; (ii) $c_S(m^*)$ is smaller; (iii) f_m is larger; (iv) α is larger.*

Therefore, our empirical prediction is that markets in which learning effect is stronger, input customization is easier, facility cost is higher, and final products are more substitutable are more likely to permit a stable outsourcing equilibrium, holding other factors constant. While the first three predictions are intuitive, the reason for the last finding is that price competition is more intense in high elasticity markets, so the returns to marginal cost reduction are larger. We will try to measure these components of the model and test our hypotheses in the next section.

3. Empirical evidence

3.1. Data

We hand-collected systematic data on 118 semiconductor firms from around the world comprising the firms listed in the 2007 Semiconductor Data Book supplemented by the addition of member firms of the Global Semiconductor Association whose 2006 sales were US$100 million or higher. We tried to extract accurate information concerning the firms and their products by referencing various sources such as the 2006 annual reports published by the firms, product data sheets, corporate brochures, as well as press releases from the firms.

The total number of products in our sample is 387. These products are differentiated in many dimensions, so the design tools and process technologies used to produce each product vary across products as well as firms. For instance, a firm can sell two logic chips that have different design costs (e.g., one that is based on the EDA tools and another that is not). We thus classify each product into one of the five main categories (which are Discretes and ICs of which there are four—Analog; Logic; Memory; Micro) and further 19 subcategories.[22]

Our dependent variable is an indicator for production outsourcing. We carefully examined whether each semiconductor product was produced in-house or outsourced based on the multiple sources mentioned above, and constructed a variable, *Outsourcing*, that takes a value of 1 for external production and 0 otherwise.[23] Note that a semiconductor firm may produce some products in-house and outsource others. Given the rapid pace of technological developments, we observe that even traditional IDMs outsource some products.

To test our hypotheses, we first classify the design methods for each product according to whether the product in question is a cell-based design or a transistor-based design. Cells are standardized function-level units that comprise combinations of transistors and have pre-selected device properties. For this reason, when the physical design method is cell-based, the level of coordination between design and manufacturing teams is relatively low (i.e., the unstructured dialog is low) and the dependence on the EDA tools would be high.

[22] According to the World Semiconductor Trade Statistics product classification system, they are Discrete (Optoelectronics, Discrete Sensor, Other Discrete); Analog (Application Specific Standard Product, Customer Specific Integrated Circuit, Standard Linear); Logic (ASSP, CSIC, Standard Logic, Programmable Logic, Display Driver); Memory (Dynamic Random Access Memory, Static RAM, Flash Memory, Read Only Memory, Electrically Erasable Programmable ROM); and Micro (Micro Processing Unit, Digital Signal Processor, Micro Controller Unit).

[23] An outsourcing decision for a given product rarely changes, so a within-product variation is not available to test our hypotheses. However, we acknowledge that there can be unobservable factors that can correlate with the product attributes and the outsourcing decision.

On the other hand, transistor-based design involves circuits and layouts at the transistor level that depend on the experience/skills of designers to optimize the design parameters. In this case, the level of communication between design and manufacturing teams is relatively high (so the asset specificity is high). Thus, we constructed a variable, *Design_Method*, that takes a value of 1 for products that use a cell-based design method and 0 otherwise. We treat this variable as a proxy for the input customization cost.

Secondly, we classify the manufacturing process as CMOS and other special processes (such as bipolar and compound). In a CMOS process, device fundamentals can be characterized by design parameters. In particular, foundries provide their customers with libraries, which contain standardized cells, parameters, and rules, so that production engineers can build multiple products with relatively little costs to re-optimize the process.[24] Thus, economies of scope/scale can be more easily achieved using a bulk CMOS process (built on silicon wafers).

In a more customized process, the degree of versatility is low and even if some steps are presumably shared with the CMOS process, there are areas in which the foundry's learning curve may not help as much as in a CMOS process.[25] Thus, if a foundry were to use a non-CMOS process, it would be for specific products, primarily analog, and the product performance would move along the learning curve much more slowly (if any). Accordingly, we constructed a variable, *Process_Technology*, that takes a value of 1 for products that use a CMOS process and 0 otherwise. This variable captures the learning effect via economies of scope/scale.

Thirdly, we gathered estimates of the acquisition or construction cost of fabrication facilities associated with each product type. Industry reports often said that the increasing cost of building a leading-edge facility was one of the primary reasons for the rise of the foundry model. We were only able to collect some estimates or ranges of the facility costs, so we constructed a variable, *Capital_Investment*, that takes a value of 0 for facilities below USD 99 million, 1 for those in the USD 100–999 million range, and 2 for those costing more than USD 1 billion. Thus, the effect of manufacturing facility costs can be positively biased.

Finally, we constructed a variable, *Application_Range*, that takes a value of 0 for custom-manufactured products, 1 for products with a specialized set of applications (e.g., automotive, mobile phone, PC, etc), and 2 for general-purpose products (e.g., memory chips). We use this variable as a rough proxy for the consumer's elasticity of substitution, so that *Application_Range* has a higher value when products are more substitutable. The logic here is that if the range of applications is large, then there are more likely to be potentially substitutable products in the market; however, we acknowledge that this is not a direct measurement.

In addition to these variables, we collected each sample firm's relevant patent stocks because patents can capture the firm's historical capabilities, which may correlate with the make-or-buy decision. Specifically, we extracted the registered US patents owned by the sample firms from 1980 to 2007 in Section G (Physics) and Section H (Electricity) of the International Patent Classification.[26] We then selected the subsections that are important to semiconductors, divided them into those concerning process or design, and aggregated the stock of these two types of patents, using a depreciation rate of ten percent per annum (e.g., Bessen, 2008).

Other standard controls include a firm size proxy (natural log of firm revenue), as smaller firms are more likely to be financially constrained, and indicators for North American and Japanese firms, respectively, as the firm's geographic location may influence the relative costs (e.g., human resources) of semiconductor design and/or production. For instance, many fabless design firms have been founded in North America, where it is relatively easy to find design expertise as well as venture capital financing, compared to Japan or other countries.

Table 1 shows the summary statistics. About 35 percent of the sample products are outsourced. Some 33 percent of the products use a cell-based design; and about 64 percent use a CMOS process. The modal product is produced in a facility that costs less than US$99 million to build (i.e., the lowest capital investment category) and is used for a set of applications (i.e., the intermediate level of product range). The sample universe is well represented across geographic locations and product categories, as well as in terms of the firm revenue and patent stocks.

3.2. Estimation results

We estimate cross-sectional logit and probit models in which the dependent variable is an indicator for outsourced production (*Outsourcing*). Our main hypotheses are tested by using the four variables that represent the theoretical counterparts in Corollary 1. Namely, we expect cell-based design methods that reduce the input customization cost η (*Design_Method*), CMOS processes that facilitate the learning effect to take place (*Process_Technology*), higher manufacturing facilities costs f_m (*Capital_Investment*), and a greater application range as a proxy for the elasticity of substitution α (*Application_Range*) will be all positively correlated with the choice of outsourcing production.

[24] For instance, an interviewee said "Although the details of CMOS can still be a little different, the basic structure of transistors is a digital function and so it is easier to have a uniform process and output characteristics from multiple suppliers including internal factories."

[25] An interviewee said "If you have a bipolar process and you want to design on it, then you have a design that is specifically created based upon that bipolar process [...] If you are dealing with bipolar or analog process, then outputs from the two factories, external vs. internal, are different."

[26] For process patents, the chosen fields are G01 (Measuing; Testing), G02 (Optics), G03 (Photography), G05 (Controlling; Regulating), and H01 (Basic electric elements). For design patents, they are G06 (Computing; Calculating; Counting), G09 (Educating; Cryptography; Display; Advertising; Seals), G10 (Musical instruments; Acoustics), G11 (Information storage), H02 (Generation, conversion, or distribution of electric power), H03 (Basic electronic circuitry), H04 (Electric communication technique), and H05 (Electric techniques not otherwise provided for).

Table 1
Summary statistics.

Variable	Mean	Std. dev.	Min	Max
Outsourcing	0.354	0.479	0	1
Design_Method	0.326	0.469	0	1
Process_Technology	0.641	0.480	0	1
Capital_Investment	0.558	0.626	0	2
Application_Range	1.152	0.534	0	2
Analog	0.315	0.465	0	1
Discrete	0.127	0.333	0	1
Logic	0.282	0.450	0	1
Memory	0.147	0.355	0	1
Micro	0.129	0.336	0	1
Process_Patents	784	1157	0	4482
Design_Patents	1023	1569	0	5761
ln(Revenue in Millions)	7.257	1.390	3.887	10.487
North American	0.483	0.500	0	1
Japanese	0.331	0.471	0	1

Sample comprises 387 semiconductor products from 118 firms in year 2007.

Given that our specification is a cross-sectional regression, we cannot rule out that our estimation results are influenced by the unobserved heterogeneity that is correlated with both the explanatory variables of interest and the dependent variable. To our knowledge, however, it is rare that a semiconductor firm changes a make-or-buy decision for a given product. If a firm does change the production sourcing decision, then it is typically for a new product line while the sourcing of existing product lines remains unchanged.

By including the above explanatory variables in addition to the main product category dummies (which is where the most of data variation come from in the prior works), we want to separately estimate the effects of these variables from those of product categories. Note that firms in our dataset have multiple product lines, and corporate-level decisions on whether to outsource each product line may be correlated within firms. Thus, we cluster our samples at the firm level, so the standard errors for products by the same firm can be correlated.

Our empirical specification is as follows:

$$\begin{aligned} Outsourcing_i = {} & \beta_0 + \beta_1 Design\_Method_i + \beta_2 Process\_Technology_i \\ & + \beta_3 Capital\_Investment_i + \beta_4 Application\_Range_i \\ & + X_i'\delta + \varepsilon_i, \end{aligned}$$

where X_i includes four product category dummies as well as firm-level control variables, and ε_i is an error term.

Table 2 shows the estimation results. The first column shows the coefficient estimates in a logit model and the third column in a probit model. We find that the first three variables, *Design_Method, Process_Technology*, and *Capital_Investment* are positive and statistically significant at the conventional level, rendering support for our prediction. The coefficient on *Application_Range* has the correct sign, but it is not statistically significant, perhaps due to the imprecise/indirect measurement of the elasticity of substitution parameter as previously mentioned.

Next, we examine the product category effects. In both specifications, it is found that analog, logic, and memory chips are statistically significantly associated with outsourcing, relative to microchips which are the omitted category. Somewhat surprisingly, this stands in contrast with the findings in the literature. For instance, Monteverde (1995), Leiblein and Miller (2003), and Macher (2006) predict that analog and memory chips are more likely to be integrated, while logic chips are more likely to be outsourced. We do not, however, think that our findings contradict the literature because here we separately estimate the effects of key properties of the products, apart from those of product category dummies.

We also examine whether accumulating more patents in the fields that are related to the manufacturing process and/or chip design would be associated with the outsourcing decision. The same (first and third) columns indicate that the coefficients on *Process_Patents* and *Design_Patents* are not statistically significant. On the other hand, the log of firm revenue, a firm size proxy, is negatively and statistically significantly related to outsourcing, which implies that large, established firms may have better access to the capital market to finance in-house production, while smaller firms are financially constrained from doing so. Our findings tell us that the capital market access is not the only significant factor.

Finally, a dummy for firms that are located in North America is statistically significantly associated with outsourcing, while a dummy for firms located in Japan is statistically significantly associated with in-house production, relative to the rest of the world. This is in fact not too surprising given that the US had more formal training and higher education in the relevant fields and the labor market for engineers was also more flexible, which encouraged a variety of development projects compatible with outsourcing, while in Japan the labor market was more rigid with permanent employment and the firm expertise was built on experience and cohesive labor force that limited the option of outsourcing (e.g., West, 2002).

The second and fourth columns show the marginal effects based on the logit and the probit model, respectively, (where the effect is for a discrete change in indicator variables). The marginal effects are consistent across the two specifications

Table 2
Determinants of outsourcing.

Variable	Logit	Margin eff.	Probit	Margin eff.
Design_Method	1.411***	0.2786***	0.8084***	0.2730***
	(0.4490)	(0.0926)	(0.2468)	(0.0854)
Process_Technology	2.376***	0.3528***	1.398***	0.3708***
	(0.5243)	(0.0710)	(0.3005)	(0.0698)
Capital_Investment	0.7571**	0.1348**	0.4230**	0.1336**
	(0.3220)	(0.0555)	(0.1794)	(0.0553)
Application_Range	0.4518	0.0804	0.2583	0.0816
	(0.3784)	(0.0657)	(0.2077)	(0.0645)
Analog	2.731***	0.5487***	1.563***	0.5308***
	(0.7027)	(0.1256)	(0.3971)	(0.1230)
Discrete	1.107	0.2346	0.6860	0.2458
	(0.8762)	(0.2069)	(0.5048)	(0.1945)
Logic	1.651***	0.3376***	0.9405***	0.3253***
	(0.3745)	(0.0821)	(0.2135)	(0.0769)
Memory	1.262**	0.2690*	0.7857**	0.2825**
	(0.6307)	(0.1504)	(0.3532)	(0.1360)
Process_Patents	−0.0006	−0.0001	−0.0003	−0.0001
	(0.0004)	(0.0001)	(0.0002)	(0.0001)
Design_Patents	0.0002	0.0000	0.0001	0.0000
	(0.0003)	(0.0001)	(0.0002)	(0.0001)
ln(Revenue in Millions)	−0.5088***	−0.0906***	−0.2949***	−0.0931***
	(0.1428)	(0.0269)	(0.0814)	(0.0269)
North American	1.294***	0.2316**	0.7771***	0.2441***
	(0.4901)	(0.0897)	(0.2839)	(0.0888)
Japanese	−1.124**	−0.1798**	−0.5856*	−0.1710**
	(0.5508)	(0.0782)	(0.3026)	(0.0806)
N	387		387	
Correct Classification %	80.88		80.88	

The dependent variable is an indicator for outsourcing. Standard errors are clustered at the firm level and reported in the parentheses. * denotes statistical significance at the 10% level, ** at the 5% level, and *** at the 1% level.

and appear economically significant. Among the first four variables, the effect of CMOS processes is the largest, followed by the cell-based design methods and the manufacturing facility costs. As for the category dummies, analog and logic chips (and to some degree memory chips) also have large effect sizes.

For an average firm, even a significant increase in firm revenue (in a log scale) would be only associated with a few percentage point decrease in the probability of outsourcing, so the capital market access story does not seem to account for a large variation in the firm's outsourcing decision. On the other hand, both the North American and the Japanese location dummies have relatively large marginal effects. This indicates that the outsourcing decision can be influenced by institutional environments, which could be of some interests to industrial policymakers.

Specifically, we believe that there are two main reasons behind this difference between North America and Japan. First, the intrafirm coordination cost could be substantially smaller in Japanese IDMs because Japanese workers tend to have more firm-specific human capital such as broader functional knowledge within the organization under the practice of lifetime employment and job rotation. Second, the adjustment cost for adapting to new technological condition might have been higher for Japanese IDMs.

For example, many Japanese IDMs had vertical *keiretsus*, where a large number of affiliated firms are traditionally organized into tiers on the group's organizational chart, ranging from raw material to distribution. Hence, a unique semiconductor ecosystem was built around major IDMs, and it has been difficult for them to consolidate manufacturing divisions or outsource production and design to overseas semiconductor manufacturers and fabless companies, not to mention that it was infeasible to outsource them to domestic competitors (Nagai and Tanabe, 2011).

4. Conclusion

We have analyzed firms' production outsourcing decision, where the learning effect due to economies of scope/scale and the standardization or modularization of design can lead to a stable outsourcing equilibrium. Using hand-collected data on 387 semiconductor products and 19 product (sub)categories, we found some empirical support for our model's predictions. Our findings are in line with the TCE literature, and it also provides a technological interpretation for the predictions in that literature. For instance, rather than relating transaction attributes to the notion of transaction costs, we analyzed a model in which technical attributes can directly affect the make-or-buy decision.

While the TCE literature does not typically consider the interrelationship among parties on the same side who do not trade each other, our equilibrium analysis takes into account this interaction through the stable bargaining solution. Further, multiple equilibria appear in our model. This suggests that in addition to individual firms making a make-or-buy decision,

equilibrium coordination among firms can be important, and accordingly an industry-wide shift from integration to outsourcing could happen once the outsourcing equilibrium satisfies the stable condition. It would be useful for future research to re-examine the role of design methods and process technologies in other industries.

Author Contribution

All authors made equal contributions to the article.

Appendix

Proof of Lemma 1. The proof is by induction on the number of supply relationships with a manufacturer, m. If the manufacturer only collaborates with one designer ($m = 1$), then the stability condition requires $S(1) = B_i(1)$, which is $T_i^*(1) - c_S(1)y_i^*(1) = p_i^*(1)y_i^*(1) - T_i^*(1)$. Rearranging yields $T_i^*(1) = c_S(1)y_i^*(1) + \frac{1}{2}V(1)$.

Next, consider the outcome with two collaborating designers. Suppose one of the designers wishes to renegotiate with the manufacturer. If the designer terminates the relationship, then the manufacturer loses $S(2) - S(1)$ on the margin. Hence, the two parties would negotiate the terms such that $S(2) - S(1) = B_i(2)$, which means after substitution $2T_i^*(2) - 2c_S(2)y_i^*(2) - \frac{1}{2}V(1) = p_i^*(2)y_i^*(2) - T_i^*(2)$. Simplifying yields $T_i^*(2) = c_S(2)y_i^*(2) + \frac{1}{6}(V(2) + V(1))$.

Thus, we conjecture that generalizing this argument for any m, the stable outcome would be characterized by

$$T_i^*(m) = c_S(m)y_i^*(m) + \sum_{k=1}^{m} \frac{V(k)}{m(m+1)}.$$

Now, suppose the number of supply relationships for the manufacturer is $m+1$. Then the stability condition requires $S(m+1) - S(m) = B_i(m+1)$. Substituting $T_i^*(m)$ and rearranging yield

$$\begin{aligned} T_i^*(m+1) &= \sum_{k=1}^{m} \frac{V(k)}{(m+1)(m+2)} + \frac{V(m+1)}{(m+1)(m+2)} + c_S(m+1)y_i^*(m+1) \\ &= \sum_{k=1}^{m+1} \frac{V(k)}{(m+1)(m+2)} + c_S(m+1)y_i^*(m+1). \end{aligned}$$

□

Proof of Proposition 1. From the above lemma, it follows that $S(m) = \sum_{k=1}^{m} \frac{V(k)}{m+1}$ and $B_i(m) = \frac{V(m)}{m} - \sum_{k=1}^{m} \frac{V(k)}{m(m+1)}$. Substituting in for (4) and (5) yields

$$\frac{1}{m+1}A\left(\frac{1}{\alpha} - 1\right)\alpha^{\frac{1}{1-\alpha}} \sum_{k=1}^{m} kc(k)^{-\frac{\alpha}{1-\alpha}} - m(1-\lambda^*)\frac{\eta}{4n^M} - f_m = 0,$$

$$A\left(\frac{1}{\alpha} - 1\right)\alpha^{\frac{1}{1-\alpha}} c(m)^{-\frac{\alpha}{1-\alpha}} - \frac{1}{m(m+1)}A\left(\frac{1}{\alpha} - 1\right)\alpha^{\frac{1}{1-\alpha}} \sum_{k=1}^{m} kc(k)^{-\frac{\alpha}{1-\alpha}} - \lambda^*\frac{\eta}{4n^M} - f_d = 0,$$

respectively. Pooling these two equations, we get

$$A_O c_S(m)^{-\frac{\alpha}{1-\alpha}}\left(\frac{1}{\alpha} - 1\right)\alpha^{\frac{1}{1-\alpha}} = \frac{\eta}{4n^M} + \frac{f_m}{m} + f_d,$$

which determines the level of market demand, A_O^*, as a function of n^M and m.

On the other hand, (6) can be rewritten as

$$A_I c_I^{-\frac{\alpha}{1-\alpha}}\left(\frac{1}{\alpha} - 1\right)\alpha^{\frac{1}{1-\alpha}} = f_m + f_d,$$

which determines the level of market demand, A_I^*, as a constant number.

Let m^* denote an equilibrium entry ratio between design firms and input suppliers and hold it fixed. Then, the number of integrated firms (n^I) and the number of design firms (n^D) that are consistent with a given level of market demand (i.e., expected profits being zero) is described by (3).

First, when $A = A_I^*$, (3) is a straight line (VV-line, as represented in Fig. 2), where the x-intercept is $n_V^D = \frac{E}{A_I^*}\left(\frac{\alpha}{c_S(m^*)}\right)^{-\alpha/(1-\alpha)}$.

Second, when $A = A_O^*(n^M)$, (3) is a nonlinear curve (OO-curve, as represented in Fig. 2), where the (non-zero) x-intercept is $n_O^D = \frac{E}{A_O^*(n^M)}\left(\frac{\alpha}{c_S(m^*)}\right)^{-\alpha/(1-\alpha)}$. This is because holding m^* constant, $A_O^*(n^M)$ is a decreasing function of n^M (hence, n^D), whereby $A_O^*(n^M) \to \infty$ as $n_O^D \to 0$; and also $n_O^I \to 0$ as $n_O^D \to 0$.

Further, it can be shown that

$$\frac{\partial n_O^I}{\partial n_O^D}|_{A=A_O^*(n^M)} = \frac{\partial n_V^I}{\partial n_V^D}|_{A=A_I^*} - \frac{E}{A_O^*(n^M)^2}\left(\frac{\partial A_O^*(n^M)}{\partial n^D}\right),$$

where the second term is positive because $\frac{\partial A_O^*(n^M)}{\partial n^D} < 0$.

This means that the slope of OO-curve is always less steep than that of VV-line. Thus, it follows that a pure outsourcing equilibrium ($n_O^D > 0, n_O^I = 0$) is stable if and only if $n_O^D > n_V^D$, or equivalently, $A_O^* < A_I^*$. □

References

Aizcorbe, A., Kortum, S., 2005. Moore's law and the semiconductor industry: a vintage model. Scand. J. Econ. 107, 603–630.
Bessen, J., 2008. The value of u.s. patents by owner and patent characteristics. Res. Policy 37, 932–945.
Coase, R., 1937. The nature of the firm. Economica 4, 386–405.
Dixit, A., Stiglitz, J., 1977. Monopolistic competition and optimum product diversity. Am. Econ. Rev. 67, 297–308.
Flamm, K., 2003. Microelectronics innovation: understanding moore's law and semiconductor price trends. Int. J. Technol. Policy Manag. 3, 127–141.
Fuller, D., Akinwande, A., Sodini, C., 2003. Leading, following or cooked goose? Innovation successes and failures in taiwan's electronics industry. Ind. Innovat. 10, 179–196.
Grossman, G., Helpman, E., 2002. Integration versus outsourcing in industry equilibrium. Q. J. Econ. 117, 85–120.
Grossman, G., Helpman, E., 2005. Outsourcing in a global economy. Rev. Econ. Stud. 72, 135–159.
Gruber, H., 1992. The learning curve in the production of semiconductor memory chips. Appl. Econ. 24, 885–894.
Gruber, H., 1998. Learning by doing and spillovers: further evidence for the semiconductor industry. Rev. Ind. Org. 13, 697–711.
Irwin, D., Klenow, P., 1994. Learning-by-doing spillovers in the semiconductor industry. J. Polit. Econ. 102, 1200–1227.
Jovanovic, B., Rousseau, P., 2002. Moore's law and learning by doing. Rev. Econ. Dyn. 5, 346–375.
Leiblein, M., Miller, D., 2003. An empirical examination of transaction- and firm-level influences on the vertical boundaries of the firm. Strateg. Manag. J. 24, 839–859.
Lieberman, M., 1991. Determinants of vertical integration: an empirical test. J. Ind. Econ. 39, 451–466.
Linden, G., Somaya, D., 2003. System-on-a-chip integration in the semiconductor industry: industry structure and firm strategies. Ind. Corp. Change 12, 545–576.
Macher, J., 2006. Technological development and the boundaries of the firm: a knowledge-based examination in semiconductor manufacturing. Manag. Sci. 52, 826–843.
Macher, J., Mowery, D., 2004. Vertical specialization and industry structure in high technology industries. Adv. Strat. Manag. 21, 317–356.
Masten, S., 1984. The organization of production: evidence from the aerospace industry. J. Law Econ. 27, 403–417.
Monteverde, K., 1995. Technical dialog as an incentive for vertical integration in the semiconductor industry. Manag. Sci. 41, 1624–1638.
Monteverde, K., Teece, D., 1982. Supplier switching costs and vertical integration in the automobile industry. Bell J. Econ. 13, 206–213.
Nagai, A., Tanabe, K., 2011. The role of semiconductor distributors in the japanese semiconductor market. In: Matsuo, T., Fujimoto, T. (Eds.), E-Activity and Intelligent Web Construction: Effects of Social Design. IGI Global, pp. 111–118.
Riordan, H., Williamson, O., 1985. Asset specificity and economic organization. Int. J. Ind. Org. 3, 365–378.
Saito, M., 2009. Global semiconductor industry trend—IDM versus foundry approaches. Proc. IEEE 97, 1658–1660.
Stole, L., Zwiebel, J., 1996. Intra-firm bargaining under non-binding contracts. Rev. Econ. Stud. 63, 375–410.
Tadelis, S., Williamson, O., 2012. Transaction cost economics. In: Gibbons, R., Roberts, J. (Eds.), The Handbook of Organizational Economics. Princeton University Press, pp. 159–190.
Waldrop, M., 2016. The chips are down for moore's law. Nature 530, 144–147.
West, J., 2002. Limits to globalization: organizational homogeneity and diversity in the semiconductor industry. Ind. Corp. Change 11, 159–188.
Williamson, O., 1975. Markets and Hierarchies: Analysis and Antitrust Implications. Free Press.
Williamson, O., 1985. The Economic Institutions of Capitalism. Free Press.