

Liu, An-Hsiang; Siebert, Ralph; Zulehner, Christine

Working Paper

The Impact of Entry Regulation on Total Welfare: A Policy Experiment

CESifo Working Paper, No. 4291

Provided in Cooperation with:

Ifo Institute – Leibniz Institute for Economic Research at the University of Munich

Suggested Citation: Liu, An-Hsiang; Siebert, Ralph; Zulehner, Christine (2013) : The Impact of Entry Regulation on Total Welfare: A Policy Experiment, CESifo Working Paper, No. 4291, Center for Economic Studies and ifo Institute (CESifo), Munich

This Version is available at:

<https://hdl.handle.net/10419/77657>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.



The Impact of Entry Regulation on Total Welfare: A Policy Experiment

An-Hsiang Liu
Ralph Siebert
Christine Zulehner

CESIFO WORKING PAPER NO. 4291
CATEGORY 11: INDUSTRIAL ORGANISATION
JUNE 2013

Presented at CESifo Area Conference on Applied Microeconomics, March 2013

An electronic version of the paper may be downloaded

- from the SSRN website: www.SSRN.com
- from the RePEc website: www.RePEc.org
- from the CESifo website: www.CESifo-group.org/wp

The Impact of Entry Regulation on Total Welfare: A Policy Experiment

Abstract

This paper evaluates how different lengths of entry regulation impact market structure and market performance using a dynamic structural model. We formulate an oligopoly model in the tradition of Ericson and Pakes (1995) and allow entry costs to vary over time. Firms have the opportunity to produce multiple products, and decide when to enter a market, followed by production and exit decisions. Using quarterly firm-level data on the static random access memory industry from 1974 to 2003, we find that entry costs decline by more than 90% within the first three years. Our policy experiments provide evidence that the duration of entry regulation has a negative impact on consumer surplus. We also find that entry protection increases total surplus if the protection duration is either sufficiently short or sufficiently long. If entry protection duration is short, the increase in monopolist's profits and entry cost saving dominate the reduction in consumer welfare, which affects total welfare positively. If protection duration is long, dynamic efficiency gains, i.e., the delay of subsequent entry and savings on entry costs impact total welfare positively.

JEL-Code: C100, L100, L600, O300.

Keywords: development costs, entry costs, dynamic oligopoly, entry regulation, innovation, intellectual property rights, market entry, market structure, policy experiment, semiconductor industry.

An-Hsiang Liu
Wuhan University Luojiashan
Wuhan / China
anhliu0512@126.com

Ralph Siebert
Purdue University
403 West State Street
USA – West Lafayette, IN 47907-2056
rsiebert@purdue.edu

Christine Zulehner
Johannes Kepler University
Altenberger Strasse 69
Austria – 4040 Linz
christine.zulehner@jku.at

April 2013

We thank seminar participants at Purdue University, Toronto University, IUPUI and the University of Mannheim for their useful comments. We especially thank Victor Aguirregabiria, Jack Barron, Lanier Benkard, Alberto Galasso, Gautam Gowrisankaran, David Hummels, Stephen Martin, Stephen Ryan, Philipp Schmidt-Dengler, Matthew Mitchell, Carlos Serrano, Anson Soderbery, Yuya Takahashi, Justin Tobias, Mihkel Tombak and Chong Xiang for their helpful discussions and valuable suggestions on this or an earlier draft. We gratefully acknowledge financial support from the Jubiläumsfonds of Oesterreichische Nationalbank (grant 11038). All errors are our own.

1 Introduction

Many policy makers and economists argue that market entry stimulates competition and economic growth. Prominent studies show that market entry lowers prices.¹ Free or unrestricted entry, however, can lead to social inefficiency, see Chamberlin (1933), Spence (1976a, 1976b), Dixit and Stiglitz (1977), Sutton (1991), and Anderson, DePalma, and Nesterov (1995). This especially applies to markets characterized by homogeneous products, imperfect competition and high fixed or entry cost. In this case, a new entrant duplicates entry costs and "steals business" from incumbents.² Since the new entrant reduces the equilibrium output of an incumbent, the fixed costs increases average costs. Consequently, the number of firms entering a market might be above the socially desirable number, which is also referred to as excessive entry, see e.g., von Weizsacker (1980), Mankiw and Whinston (1986), and Perry (1984).

Only few empirical studies exist in evaluating excessive entry, and most studies assume that the fixed entry costs remain stable over time. Excellent studies in this area are Berry and Waldfogel (1999), who focus on the U.S. commercial radio broadcasting industry and show that free entry can lead to social inefficiency. They predicted that the welfare loss due to free entry amounts to 45% of firms' revenue. Hsieh and Moretti (2003) show that a socially inefficient large number of real estate agents enter markets characterized by high housing prices. Many practitioners and managers, however, highlighted the fact that entry costs often strongly decline over time. One example are high-tech industries, as the cost of capital, such as investments into new manufacturing equipment and new plants depreciates quickly.³ In industries characterized by drastically declining entry costs, the problem of excessive entry becomes more severe. From a social point of view firms pay too high of an entry costs from entering too early. In this case, entry restrictions or regulation can often be socially desirable.

Up to date very few studies that focus on excessive entry costs and entry regulation. The aim of our study is to empirically investigate the long-run impact of entry restrictions or regulation on consumer and producer surplus applying a fully dynamic structural model. This paper explicitly focuses on the role of excessive and declining entry costs. We formulate a dynamic

¹Excellent examples in this area are Berry (1992), Bresnahan and Reiss (1990, 1991), Campbell and Hopenhayn (2005), Carlton (1983), Davis (2006), Dunne, Roberts and Samuelson (1988), Geroski (1989), Klepper (2002), Mazzeo (2002), Seim (2006), and Toivanen and Waterson (2000 and 2005).

²A private firm does not take into account that increases in its own output cannibalizes the output of other firms. Firms consider only their private benefits, but not the social returns. A social planner, however, would incorporate this effect.

³For example, a new semiconductor manufacturing plant can grow obsolescent in only 3 years.

oligopoly model in the tradition of Ericson and Pakes (1995), in which firms make entry, exit and production decisions. Firms are assumed to be rational and forward looking agents who maximize their discounted payoffs. Since entry costs are allowed to decline over time, firms face a trade off in determining the optimal time to enter: An early market entry requires a higher entry cost, but ensures higher profit streams due to facing fewer firms and lower competition.⁴ A late market entry saves on entry costs, but lowers firms' discounted profit streams. Late entry returns lower discounted profits, but requires firms to pay lower entry costs.

We use data from the Static Random Access Memory (SRAM) industry which belongs to the family of semiconductors and which are used as an input for electronic devices.⁵ This industry provides a natural setting for our study for several reasons. First, firms enter new markets at different time periods. For example, more than 50% of the firms still entered the memory industry 5 years after the first firm had entered, which suggests that late entry still is a profitable option and entry costs decline over time. Second, SRAM chips are classified according to different generations depending on the performance and information storage capacity of memory chips. The existence of different generations offers a rich set of entry observations. Moreover, firms enter new product generations at different points of time throughout the life cycle, which allows the identification of time-varying entry costs. Third, the SRAM industry is characterized by free market entry. For example, firms are not able to patent the information storage capacity of those memory chips.⁶ Hence, this industry represents an appropriate reference case to perform our policy experiments, i.e., to evaluate the impact of different lengths of entry protection on total surplus. Fourth, industry experts, managers and engineers frequently refer to the fact that the SRAM industry is characterized by a high depreciation of capital and machinery which cause the decline in entry costs. As mentioned above, the entry costs determined by the progress of the production machines which are developed outside the SRAM industry. Hence, entry cost can be considered exogenous. Moore's law predicts a duplication of storage capacity and with that an emergence of a new product generation every two to three years. In such an industry, in contrast to, for example, the pharmaceutical industry, strategic waiting does not play a role.

To estimate our model, we apply a two-stage estimation procedure by Bajari, Benkard and

⁴See e.g., Fudenberg and Tirole (1985), Riordan (1992), and Klette and Kortum (2004).

⁵A detailed industry description is provided in Section 2.

⁶The patents rather protect process innovation and manufacturing processes which improve production costs, rather than new technologies that are related to the introduction of new memory chips. Section 2.1 explains the facts in more detail.

Levin (2007). The first stage estimates firms' entry, exit and production policies. Based on the policy functions, we perform forward simulations and calculate the discounted continuation values. In a second stage, we perturbate the entry times to identify the evolution of entry costs over time. Using quarterly firm-level data from 1974 to 2003, our estimation results show reasonable estimates for the policy functions and continuation values. The results also describe a considerable decline in entry costs of more than 90% within the first three years. Based on the structural estimates, we perform policy experiments and evaluate the long-run impact of how entry regulations of different lengths impact market structure, performance and consumer and producer surplus. The results from the policy experiment show that a longer entry regulation continuously reduces consumer welfare compared to the free entry case. Most interestingly, total welfare increases if the length of entry regulation is sufficiently short (less than 1 year) or long (more than 6 years). If entry regulation is short, entry cost savings cover the losses in consumer surplus and industry profits, and total welfare increases. If the length of entry regulation is in the intermediate range, the increase in producer surplus and the savings on entry costs from prohibiting excessive entry is not large enough to compensate for the loss in consumer surplus. If the entry regulation is sufficiently long, excessive entry is limited which saves on entry costs. Moreover, the monopolist's profits significantly increase due to own learning effect. Both effects, i.e., the entry cost savings and the monopolist's gain in profit, dominate the losses incurred by prohibiting other firms to enter the market. Consequently, total welfare increases.

The remainder of the paper is organized as follows. Section 2 describes the institutional features of the semiconductor industry and present summary statistics. Section 3 introduces our dynamic oligopoly model and Section 4 describes the econometric model. In Section 5 we discuss the estimation results and the results for our counterfactual experiments. We conclude in Section 6.

2 Institutional and industry background

In this section, we briefly describe the institutional background of the semiconductor industry. Then, we describe the SRAM industry in more detail and also provide a description of the data and summary statistics.

2.1 Entry regulations

Entry regulations are imposed by governments and can take on various forms such as patents, entry licenses etc. Licenses are frequently used to avoid excessive entry in industries that are characterized by high entry costs like the telecommunication or semiconductor industry. The purpose of limiting access to markets is that the associated fixed or entry cost is high, and to avoid excessive entry. Entry licenses also serve the purpose to protect early entrants from successive entrants for a specific period of time. Entry into new markets frequently requires high investments which is considered as an entry impediment. Protecting incumbents from successive entry for a limited period of time also allows them to earn higher profits and to sooner recover their entry costs. The provision of intellectual property rights in form of patents is one common alternative to regulate entry.⁷ Patents protect innovators from successive entries by granting innovators a monopoly position for a certain time period.⁸ It should be noted at this point that in the semiconductor industry, firms mostly file patents for new process technologies rather than new products, i.e., firms are not able to claim a patent on the capacity of information storage on memory chips. In fact, semiconductor managers mentioned that almost all patents granted refer to new manufacturing processes, rather than new generations or the associated memory capacity, chip and architectural designs. Hence, firms' entry costs mostly refer to the cost of capital and the investment in production machineries.

Several policy debates discuss the advantages and disadvantages of entry regulation. One frequent concern is that entry regulations result in temporary higher prices causing harm to consumer surplus. For example, Gilbert and Shapiro (1990) remark that the negative impact on consumer welfare make limitations on the length of entry protection necessary. In fact, the length of entry protection varies substantially across countries and has frequently been modified over time, see Moser (2005).

Chaudhuri, Goldberg and Jia (2006) empirically evaluates the short-run impact of patent lengths on welfare. They find that the Indian pharmaceuticals market would suffer from enormous losses in consumer welfare if the Indian pharmaceutical market was protected by patent laws similar to the U.S. Hughes, Moore and Snyder (2002) find that elimination of patent rights in pharmaceuticals yields large benefits to existing consumers. Other studies, which simulate

⁷See also Nordhaus (1969) and Lerner (2002) for further information on patent policies.

⁸In order to receive a patent, the inventor must prove that the invention is useful, novel and not obvious.

the welfare losses implied by patent protection, are e.g., Fink (2000), Maskus and Konan (1994), Subramanian (1995), and Wattal (1999 and 2000).

2.2 The SRAM industry

In the following, we describe the innovation and manufacturing process for SRAMs, which belong to the family of semiconductors. SRAMs are primarily used as an input in computers, electronic devices like digital cameras or cell phones, automotive products and household appliances, among many others. The semiconductor industry is a high-tech industry and has a significant impact on productivity growth. Most firms in this industry invest more than 30% of their revenues in R&D.

Semiconductors consist of memory chips, microprocessors, and application-specific integrated circuits. SRAMs are memory chips designed for the purpose of storing and retrieving information. The chips are classified into generations according to their capacities of information storage. A higher capacity of SRAM chips requires advanced technological improvements along different dimensions, i.e., different architectures of transistor cells which allow for smaller transistor and cell sizes, an increase in the die area, lower temperature for performing operations, lower energy consumption, advanced masks and lenses to improve etching processes. The most critical issue in the advancement of semiconductor technology is the feature size. The next generation memory chips often require finer feature sizes. For instance, the minimum feature size for 256K DRAM chips was 1.6 microns, while the same figures for the latter 1M and 4M DRAMs had been improved to 1.2 and 0.8 microns. In addition, each new generation also require one or two more mask layers. [These new manufacturing requirements drive demand for new equipment or new fabs. For example, in 2004, Samsung, Hynix, Toshiba, Intel and Micron announced a capital spending between 2 and 4.5 billion U.S. dollars. Note that these numbers are rarely reported and also represent averages only.] It should be noted that firms introduce generations successively, and do not skip one generation to achieve an early headstart on the successive generation.

In the SRAM industry, the cost of entering a new generation of SRAM chips includes the initial purchase price and installation of new fabs and equipment for producing the higher capacity memory chips. Interviews with experts and managers in the industry reassured the fact, that cost of capital is the main determinant of entry and entry cost primarily consist of the

costs of equipment. One example is the semiconductor industry, which requires high investment into new plants equipped with specific production machines. The costs of equipment, which are developed outside the semiconductor industry, significantly decline over time. Up to date, however, very little is known on how entry costs evolve over time. This is mostly due to the fact that the cost of entry is rarely observed by econometricians.⁹

Firms must constantly renovate and rebuild manufacturing capacity. In the semiconductor industry, the equipment depreciation is roughly half the cost of producing a wafer. The technological progress of memory chips is frequently described by Moore’s Law. According to Moore’s Law, the number of transistors per chip doubles every two years. A consequence of Moore’s Law is that the number of transistors results in a fourfold increase in memory capacity per chip which is considered as a new generation of memory chips. As shown in Figure 3, the introductions of the next generations occurs every 3 years consistently in spite of the facts the market structures or the patterns of entries differ significantly across generations. The accuracy of Moore’s prediction confirms the fact that the technological progress is a function of time. To summarize, the timing of entry determines the entry cost and investments that has to be incurred.

The manufacturing process of SRAMs is highly complex. The production process of every new chip generation is characterized by significant learning effects throughout the life cycle. Advanced photolithographic and chemical processes to etch electrical circuits onto the wafer surface are necessary to improve a chip’s performance. Memory chips are cut from silicon wafers, which determines silicon as the base material entering the manufacturing process. The production yield rate, which is defined as the percentage of wafers that successfully pass all production stages, can increase from around 20% to more than 80% over the life cycle, due to learning effects.

2.3 Data description and summary statistics

Next, we introduce the data sources and provide some descriptive statistics. Our study uses firm level and industry level information on prices and quantities for different SRAM generations which are compiled by Gartner Inc. The data set encompasses six product generations, namely the 4Kb, 16Kb, 64Kb, 256Kb, 1Mb and 4Mb generations. It entails quarterly data from January

⁹Several studies related to market entry focus on the timing of adopting new technologies, see e.g., Akerberg and Gowrisankaran (2006), Einav (2010), Genesove (1999), Gowrisankaran and Stavins (2004), Schmidt-Dengler (2006) and Sweeting (2006 and 2009).

1974 until December 2003. The data set covers firm and industry units shipped, the average selling price, and the number of firms in the market.

Figure 2 shows the average prices of the 4Kb to 4Mb SRAM generations over time. Prices are rapidly declining throughout the product life cycle. An important phenomenon that explains this price decline for SRAMs are learning effects, see the discussion above. Figure 1 shows how industry shipments of the 4Kb to 4Mb SRAM generations evolve over time. Shipments of each generation are characterized by product life cycles, which take on similar shapes across different generations. At a certain point of time, multiple SRAM generations are offered on the market.

Figure 3 shows the number of firms present in the 4Kb to 4Mb SRAM generations over time. The number of firms follows an inverse U-shape within every generation. Interestingly, firms continue to enter the market many years after the product generation has been launched. To illustrate market structure in more detail, Figure 4 shows the entry process for the 64Kb generation. The figure illustrates that firms enter the 64Kb generation even 20 years after the generation has been launched. We observe that about half of the firms enter in the first 5 years. Most firms enter during the 8th to the 33rd quarter while less than 10% of the firms enter in the first two years.

The fact that firms simultaneously produce multiple product generations indicates that firms are able to benefit from learning effects across generations. In light of intergenerational learning effects, multiproduct firms might have an incentive to produce more output in previous generations to gain more experience and lower their production costs in the new generation, which increases profits and allows them to enter markets earlier. It is important to mention that intergenerational learning is one major reason why (the heterogeneous) firms successively enter into different generations - rather than skipping one generation - in order to achieve an early headstart for production in subsequent generations. This fact is important in our policy experiment to make sure that firms will not skip producing a generation once entry protection become sufficiently long. Moreover, multiproduct firms also account for substitution patterns originated from the demand side. If chips are substitutes, the production of different chip generations will cannibalize the sales of other generations. Consequently, multiproduct firms will account for these substitution patterns when setting their production schedules for different generations.

3 Dynamic oligopoly model

In this section we construct a discrete-time infinite horizon model with time indexed by $t = 0, 1, \dots, \infty$. Firms, denoted by $i = 1, 2, \dots, N$, maximize the sum of profits over all periods for a specific generation k .¹⁰ The model is formulated as a state space game in which firms use Markov strategies. Firms' actions in a given period determine not only their own and rival firms' current profits, but also their own and rival firms' future states. We build on the fact that firms are rational and forward looking, i.e. they derive their discounted profit streams given the evolution of the state vector and their actions. At the beginning of each period, firms simultaneously decide whether to enter generation k and how much to produce in generation k . Firms who entered in generation k in the past, also decide whether to exit generation k or not.¹¹ We focus on pure entry strategies. Entering a new product generation incurs an entry cost of C_t^k , where the subscript t refers to the time period when entry occurs. The entry cost is deterministic and larger than zero. Since the entry cost is primarily a function of depreciation of investments, we assume the function to be industry and generation specific, rather than firm-specific.¹²

The observable state variables to every firm at period t are denoted by a vector $s_{it} = (n_t^k, x_{it}^k, x_{-it}^k, C_t^k, z_t^k, X_{it}^{k-1})$, which contains all of the payoff relevant state variables for generation k . The variables n_t^k refer to the number of firms in period t for generations k and are supposed to capture the degree of competitiveness in the markets for generation k . We also account for firms' learning capabilities. Firms are supposed to learn from past experience lowering marginal costs. We distinguish between learning experience accumulated within a generation (intragenerational learning) and learning experience accumulated in the previous generation (intergenerational learning). With regard to intragenerational learning effects, we allow firms to learn from their own experience (x_{it}^k) as well as from other firms' experience via spillovers (x_{-it}^k). z_t^k includes all the demand and cost shifters which are known to the industry.

¹⁰We simplify our entry model to one generations and believe that this is an appropriate assumption when estimating time-variant entry costs for the following reasons. First, different product generations are usually produced at different fabs. Therefore, firms' output decisions are usually made at the fab level. Second, firms that consider to enter generation k in period t only know about the existence of generation $k-1$, but do not know of generation $k+1$ yet. Moreover, firms who actively produce one generation, introduce subsequent generations because of intergenerational learning effects. Therefore, firms skipping one generation to achieve an early headstart for the next generation is not a concern in the semiconductor industry. When considering the entry of generations k , we treat the states from the previous generation as predetermined and given from the data. The assumption might be restrictive, however, for other industries and should be reconsidered.

¹¹Exit does not involve any cost or scrap value.

¹²Note that firms might have different entry times in generation k due to the fact that we allow for heterogeneous firm sizes in generation $k - 1$.

$X_{it}^{k-1} = (n_t^{k-1}, x_{it}^{k-1}, x_{-it}^{k-1})$ are state variables from the previous generation. These variables are assumed to be exogenous and given from the data. The entry of generation k may be determined by the relative competitiveness in the markets of generations $k(n_t^k)$ and $k-1(n_t^{k-1})$. We allow firms to learn from their own production experience in the previous generation (x_{it}^{k-1}). We also account for firms' intergenerational learning effects via spillovers resulting from other firms' production experience in the previous generation (x_{-it}^{k-1}). Finally, the time-variant entry cost C_t^k enter the state vector.

3.1 Evolution of states

Number of firms

We define the transition of the number of firms n_t^k within generation k as follows:

$$n_t^k = n_{t-1}^k + ne_t^k - nx_t^k, \quad (1)$$

where ne_t^k is the number of entering firms and nx_t^k the number of exiting firms.

Learning

Firm i 's marginal cost of production is characterized by several learning components. The first component describes a firm's intragenerational learning from own experience. Firm's own learning within generation k is given as follows:¹³

$$x_{it}^k = x_{it-1}^k + q_{it-1}^k, \quad (2)$$

where q_{it-1}^k is firm i 's output in generation k at time $t-1$ and $x_{i0}^k = 0$, assuming that there is no own experience in the beginning of the product cycle of each product generation. The other components describe the extent to which firm i learns from other firms' experience via spillovers within and between generations. Intragenerational spillovers are specified as follows:

$$x_{-it}^k = x_{-it-1}^k + \sum_{j \neq i} q_{jt-1}^k, \quad (3)$$

with $x_{-i0}^k = 0$, assuming that there are no spillovers in the beginning of the product cycle of

¹³Note that the laws of motion for the state variables are deterministic.

each product generation. In general, potential entrants in a specific generation may benefit from own experience in the previous generation as well as intragenerational and intergenerational spillovers.

Productivity shocks

In each period, each incumbent in generation k receives a private productivity shock ν_{it}^k . The shocks are drawn from a known normal distribution G^k with zero means and constant variances σ_k^2 and are independently and identically distributed across firms and across periods.

Marginal costs

The marginal cost function depends on the set of payoff-relevant state variables and follows the specification in Irwin and Klenow (1994) to account for own learning and spillover learning. Firm i 's marginal cost mc_{it}^k is specified as follows:

$$\ln(mc_{it}^k) = \rho_0^k + \rho_i + \rho_1 \ln(P_t^{sil}) + \rho_2 \ln \left(x_{it}^k + \rho_3 x_{-it}^k + \rho_4 \left[x_{it}^{k-1} + \rho_3 x_{-it}^{k-1} \right] \right) + \nu_{it}^k, \quad (4)$$

Marginal costs are described by generation-specific intercepts (ρ_0^k), firm fixed effects (ρ_i), the price of silicon P_t^{sil} and several learning components, i.e., intragenerational own learning (x_{it}^k) and learning from others via spillovers (x_{-it}^k), intergenerational own learning (x_{it}^{k-1}) and learning through spillovers x_{-it}^{k-1} . The parameters ρ_2 and ρ_4 measure the impact of the own intra- and intergenerational learning effects on the marginal costs. The parameter ρ_3 allows for a separate impact of own learning and learning via spillovers on marginal costs.

Strategies

Firm i 's strategy is given by $\sigma_i^k(s, \nu_i)$, where σ_i is the set of firm i 's strategies that describe the production, entry and exit decisions. The set of state vectors (s_{it}) are summarized in s and ν_i includes the generation-specific firm-level productivities. Rival firms' strategies are denoted by σ_{-i} .

3.2 Profit

Each firm maximizes its future discounted profits from generation k . If a firm is active in generation k , its per period profit is given by:

$$\pi_{it}^k(\sigma_i^k, \sigma_{-i}^k, s_t, \nu_{it}^k) = q_{it}^k[p_t^k - mc_{it}^k(\nu_{it}^k)],$$

where the variable p_t^k represents the price for generation k . The price is a function of the industry output $q_t^k = \sum_i q_{it}^k$ and observable demand shifters d_t . Hence, firm i 's per period profits for generation k at time t depend on the actions of all firms, the state vector, and firm i 's private shock. A firm earns zero profit if it is inactive in generation k .

If behavior is given by a Markov perfect equilibrium strategy profile $\sigma^k = (\sigma_i^k, \sigma_{-i}^k)$, firm i 's value function can be written recursively:

$$V_i(s, \sigma^k; \theta) = E_\nu \left[\pi_i^k(\sigma^k, s, \nu_i^k) - I[C^k](\sigma^k) + \beta \int V_i(s'; \sigma^k) dP(s' | \sigma^k, s) \middle| s \right], \quad (5)$$

where θ denotes a vector of parameters and I refers to an indicator function which takes on a value of 1 if a firm enters a market and 0 otherwise. P describes the evolution from today's state s to tomorrow's state s' .

3.3 Timing

The timing of our model can be summarized as follows. In each period t , events occur in the following order:

1. Firm i considers the following state $s_{it} = (n_t^k, x_{it}^k, x_{-it}^k, C_t^k, z_t^k, X_{it}^{k-1})$.
2. Given the state vector, potential entrants make their entry decisions for generation k and incumbents make their exit decisions for generation k .
3. Incumbent firms in generation k observe their productivity shocks ν_{it}^k , choose outputs q_{it}^k and realize profits.
4. The state moves to $(n_{t+1}^k, x_{it+1}^k, x_{-it+1}^k, C_{t+1}^k, z_{t+1}^k, X_{it+1}^{k-1})$.

4 Econometric model

In this section, we describe our econometric model. We apply the two-step estimation method developed by Bajari, Benkard and Levin (2007) to estimate the cost of entering the 64K SRAM generation. The first stage includes the estimation of the demand, the policy functions and the marginal cost functions. All policies are functions of the observed state, s , including the entry cost C_t^k and the productivity shocks ν_{it}^k which are unobserved to the econometrician. We include a time trend in all of the policy functions to proxy for the declining entry costs over time. In the second stage, we apply forward simulations to calculate the continuation values. Finally, we recover the entry cost over time that rationalizes firms' policies.

4.1 First Stage

Demand

We assume that every generation is homogeneous in itself, but different generations represent differentiated goods. The demand function is specified log-linearly as follows:

$$\ln(q_t^k) = \alpha_0^k + \alpha_1^k \ln(P_t^k) + \alpha_2 \ln(P_t^S) + \alpha_3 \ln(GGDP_t) + \alpha_4 \ln P_t^{DRAM} + \alpha_5 Time_t + d_t^k, \quad (6)$$

where we denote the vector of coefficients with α . We account for generation-specific demand shifters (α_0^k). P_t^k is the price for generation k . Note that we allow for generation-specific price elasticities (α_1^k). P_t^S is the price for the substitutable generation.¹⁴ $GGDP_t$ is the growth rate of GDP, P_t^{DRAM} is the price of DRAMs and $Time$ refers to a time trend.

Marginal Costs

In line with our theoretical model, we retrieve firms' marginal costs for the $k = 64K$ generations (mc_{it}^{64K}) using the estimated price elasticities $\hat{\alpha}_1^{64K}$ of our demand equation (6) and the observed market shares, ms_{it}^{64K} , see also Siebert and Zulehner (2011):

$$mc_{it}^{64K} = P_t^{64K} \left(1 + \frac{ms_{it}^{64K}}{\hat{\alpha}_1^{64K}} \right). \quad (7)$$

¹⁴As substitute we use the average SRAM price of the adjacent generations.

Output Policy

The output policy depends on the set of observed state variables, the productivity shock and a time trend. Hence, the output policy accounts for factor prices, demand shifters, market structure, intra- and intergenerational own learning and learning via spillovers. The output function for the 64K generation is specified as:

$$\begin{aligned} \ln(q_{it}^{64K}) = & \gamma_i + \gamma_1 \ln(P_t^{sil}) + \gamma_2 \ln P_t^{DRAM} + \gamma_3 GDP_t + \gamma_4 \ln(n_t^{16K}) + \gamma_5 \ln(n_t^{64K}) + \gamma_6 \ln(x_{it}^{16K}) \\ & + \gamma_7 \ln(x_{it}^{64K}) + \gamma_8 \ln(x_{-it}^{16K}) + \gamma_9 \ln(x_{-it}^{64K}) + \gamma_{10} Time_t + \gamma_{11} \hat{\nu}_{it}^{64K} + \epsilon_{it}^{64K}, \end{aligned} \quad (8)$$

where γ_i is a firm level fixed effect. We assume that the error term ϵ_{it}^{64K} is i.i.d normally distributed. $\hat{\nu}_{it}^{64K}$ is the residual from the marginal cost estimation and is interpreted as the productivity shock. A positive shock rises the marginal cost of production and hence reduces a firm's output. Thus, we should expect a negative sign for γ_{11} .

Entry and Exit Policies

We estimate the following probit model for entering the 64K generation:¹⁵

$$Pr(e_{it}^{64K} = 1; s) = \Phi\left(\lambda_i + \lambda_1 \ln(n_t^{64K}) + \lambda_2 \ln(x_{it}^{16K}) + \lambda_3 \ln(x_{-it}^{64K}) + \lambda_4 Time_t\right), \quad (9)$$

where $\Phi(\cdot)$ is the cumulative distribution function of the standard normal. The entry policy depends on the number of firms active in the 64K generation (n_t^{64K}), firm i 's experience from the previous 16K generation (x_{it}^{16K}), the spillover from other firms in the 64K generation (x_{-it}^{64K}) and a time trend.

The exit policies for the 64K generation is formulated as:

$$Pr(ex_{it}^{64K} = 1; s) = \Phi\left(\psi_0 + \psi_1 \ln(n_t^{64K}) + \psi_2 \ln(x_{it}^{64K}) + \psi_3 \ln(x_{-it}^{64K}) + \psi_4 Time_t\right), \quad (10)$$

where the probability of exit is a function of the number of firms, the own experience and other firms' experience in the current generation.

¹⁵Note, since our theoretical model explicitly formulates entry in generation k , we only have to estimate the entry policy for the 64K generation. This is also sufficient for our policy experiments to evaluate the lengths of entry regulations in the 64K generation.

4.2 Second Stage

Using the estimated policy functions, we take draws for the productivity shocks for each incumbent in each period and forward simulate the sequences of optimal outputs for each firm. In each simulation l , we record t_{il}^* (firm i 's optimal timing of entering the $64K$ generation in simulation l). Using $N \times L$ observations, where N is the number of firms and L is the number of simulations, we generate further potential equilibrium paths. Hence, based on the optimal timing (t_{il}^*), the demand equation, the marginal cost function, the drawn productivity shocks and the optimal policies, we compute a simulated sequence of outputs. Next, we compute the discounted profits before subtracting the entry costs, i.e., $\tilde{V}_{il}(t_{il}^*)$. To summarize, we compute the value functions applying the following steps for each simulation:

1. We draw one productivity shock for each incumbent at each period i.e., ν_{it}^{64K} .
2. We use the optimal production, entry and exit policies to update firms' production choices \hat{q}_{it}^{64K} and state variables \hat{s}_{it} . Note that the draws of productivity shocks enter \hat{q}_{it}^{64K} and \hat{s}_{it} (see equation 8).
3. For every period, we update prices $\hat{p}^{64K}(\hat{q}_{it}^{64K}, \hat{q}_{-it}^{64K})$ using the estimated demand equation.
4. We compute the vector of marginal costs $\hat{m}c_i^{64K}$ using the estimated marginal cost function, \hat{q}_{it}^{64K} , \hat{s}_{it} and ν_{it}^{64K} .
5. We calculate the value functions: $\tilde{V}_{il}(t_{il}^*, t_{-il}^*) = \sum_{t=0}^{\infty} \beta^t (\hat{\pi}_{it}^{64K})$.

In order to identify the evolution of entry costs over time, we compare the changes in the discounted values at the optimal entry times with the values at the distorted and therefore suboptimal entry times to derive the upper and lower bounds for the entry costs.

Regarding the upper bound of the entry costs, we compare the discounted values (net of entry costs) at the optimal entry time $\tilde{V}_{il}(t_{il}^*) - C(t_{il}^*)$ with the discounted value (net of entry costs) from entering one period later, i.e., $\tilde{V}_{il}(t_{il}^* + 1) - C(t_{il}^* + 1)$. Based on firms' rational expectations, we can write:

$$\tilde{V}_{il}(t_{il}^*) - C(t_{il}^*) \geq \tilde{V}_{il}(t_{il}^* + 1) - C(t_{il}^* + 1)$$

Rearranging this inequality yields:

$$\Delta C^{max} = \tilde{V}_{il}(t_{il}^*) - \tilde{V}_{il}(t_{il}^* + 1) \geq C(t_{il}^*) - C(t_{il}^* + 1), \quad (11)$$

where ΔC^{max} defines the maximum change in entry costs between period t and period $t + 1$ and represents the upper bound of the slope of the entry costs, see Figure 5.

Regarding the lower bound of the entry costs, we compare the discounted values (net of entry costs) when the optimal entry time is $(t_{il} + 1)^*$ with the discounted value (net of entry costs) evaluated at an earlier suboptimal entry time (t_{il}) :

$$\tilde{V}_{il}(t_{il} + 1)^* - C(t_{il} + 1)^* \geq \tilde{V}_{il}(t_{il}) - C(t_{il})$$

Again, rearranging the inequality gives:

$$\Delta C^{min} = \tilde{V}_{il}(t_{il}) - \tilde{V}_{il}(t_{il} + 1)^* \leq C(t_{il}) - C(t_{il} + 1)^*, \quad (12)$$

where ΔC^{min} represents the lower bound of the slope of entry cost and defines the minimum entry costs between period t and period $t + 1$, see Figure 6.

Next, we search over the entry costs that best rationalize the two conditions - inequalities (11) and (12) - using a minimum distance approach. We therefore define a penalty, when inequalities (11) and (12) are violated.¹⁶ Then, ΔC can be recovered by searching over c , that minimizes the weighted sum of penalties. Denoting $A_t \equiv \{(i, l) \mid t_{il}^* = t\}$ as the set of observations where firms optimally enter at period t , and $A_{t+1} \equiv \{(i, l) \mid (t_{il} + 1)^* = t + 1\}$ as the set of observations where firms optimally enter at period $t + 1$:

$$\widehat{\Delta C(t)} = \arg \min_c \left\{ \frac{1}{|A_t|} \sum_{A_t} \min \left[0, \tilde{V}_{il}(t_{il}^*) - (\tilde{V}_{il}(t_{il}^* + 1) + c) \right] + \frac{1}{|A_{t+1}|} \sum_{A_{t+1}} \min \left[0, \tilde{V}_{il}(t_{il} + 1)^* - (\tilde{V}_{il}(t_{il}) - c) \right] \right\} \quad (13)$$

where $|\cdot|$ refers to the number of elements in the set.

To summarize, we perform the following steps in our estimation algorithm:

¹⁶A penalty is defined when the simulated discounted value (net of entry costs) is higher at a suboptimal entry time than at the optimal entry time.

1. We estimate the demand equation (6) to retrieve the elasticity of demand.
2. We use the elasticity of demand in order to calculate the marginal costs using the Lerner Index, see equation (7).
3. We estimate the marginal cost function to retrieve the cost parameters entering in equation (4).
4. We estimate the production policies (equation 8), the entry policy (equation 9) and the exit policy (equation 10).
5. We use the cost parameters and the optimal policies to apply forward simulations and to calculate the value functions (before entry costs), i.e., $\tilde{V}_{il}(t_{il}^*)$ and $\tilde{V}_{il}(t_{il} + 1)^*$.
6. We distort the optimal time to enter and use the policies to calculate value functions (before entry costs) i.e., $\tilde{V}_{il}(t_{il}^* + 1)$ and $\tilde{V}_{il}(t_{il})$.
7. Finally, we use the penalty conditions and apply a minimum distance procedure to recover the slope of the entry costs (equation 13).

5 Estimation results and policy experiment

In this section, we describe our estimation results. We first discuss the results for the demand and the marginal cost function. Afterwards, we discuss the estimation results for the firms' production, entry and exit policies. Finally, we describe the results for the entry costs and discuss the results of our policy experiment.

5.1 First stage results

Demand

To estimate industry demand, we pool the data from all product generations and estimate equation (6) using ordinary least squares and instrumental variable estimation techniques. We instrument for potential correlation of the price with the error term using supply-side variables: price of silicon (P_t^{sil}) and cumulative industry output ($\sum_{t'=1}^{t-1} q_{t'}^k$) of the current generation. The estimation results are shown in Table 1. We estimate two specifications: the first specification incorporates a constant price elasticity across product generations (Columns 1 and 2). The

second specification incorporates generation-specific price elasticities (Columns 3 and 4). We employ a Chow test to test for identical price elasticities across product generations. The F-statistics returns a value of 6.01 which leads us to reject the null hypothesis that prices elasticities are equal to each other.

Comparing the results in columns 3 and 4, the instrumental variable estimation returns more elastic price elasticities of demand than the ordinary least squares estimation. This result indicates an upward bias of the ordinary least squares estimates and gives rise to the fact that prices are endogenous. To test for endogeneity, we conduct a Durbin-Wu-Hausman test. The test statistics is $F_{(1,300)} = 619.99$ for the first specification and it is $F_{(4,293)} = 6.40$ for the second specification. We therefore reject the exogeneity of prices in both specifications.

In the following, we discuss the results shown in column 4. The estimated price elasticities range from -2.75 to -3.19 , confirming a highly elastic market demand in the SRAM industry. In comparing the price elasticities throughout different generations, we find that the $64K$ generation is characterized by the least elastic demand, which explains the fact that the average prices in this generation stay at higher levels. Demand becomes more elastic in the subsequent $256K$ and $1M$ generations. Finally, our estimations return positive cross-price elasticities, confirming the fact that SRAM chips are substitutes across generations.

Marginal cost The marginal cost function, equation (4), is estimated using a non-linear least square procedure. We assume that the error in the marginal cost function is normally distributed. The results are shown in Table 2. Column 1 reports the estimation results without a time trend. All coefficients are significantly different from zero, carry the expected sign and reveal economically sensible magnitudes. In addition, the model fits remarkably well with an adjusted R^2 of about 0.90. We find a positive effect of input prices on output prices. The results further return a learning parameter of -0.34 , which corresponds to a learning rate of 21%.¹⁷ This result confirms that the SRAM industry is characterized by strong intragenerational learning effects. Moreover, we find that there are spillovers within one generation and across generations. The spillover effect within one generation is of about 30% (or 6% overall), i.e., firms are able to benefit from about a third of other firms' output. Compared to this effect, the intergenerational spillover effect is small. About 10% (or 2%) from the output produced in the

¹⁷The learning rate is defined as $1 - 2^{\rho_2}$, i.e. the percentage reduction of marginal cost when (past) experience doubles, see also Irwin and Klenow (1994).

earlier generation contributes to learning.

Column 2 reports the estimation results from a second specification that includes a linear time trend. The negative significant time trend reflects the fact that marginal costs decline over time. The inclusion of the time trend slightly alleviates the magnitude of intra- and intergenerational learning. The learning effects, however, still remain economically and statistically significant.

Since the residual of the marginal cost estimation is interpreted as private productivity shocks, we use the standard deviation of the residuals, 0.3, for the distribution of productivity shocks. Therefore, we take draws of productivity shocks from a normal distribution with zero mean and 0.3 standard deviation in the forward simulation.

Output policy

We estimate the output policies specified in equation (8). Table 3 shows the estimation results for the 64K generation using an ordinary least square estimator.¹⁸ An increase in the price of silicon increases marginal costs and hence reduces output. The coefficients of the demand shifters carry the expected signs. Firms produce more when the size of the market is larger (GDP) and when the price of a substitute is higher (P^{DRAM}). The coefficients of the number of firms within a generation is negative and significant. A more competitive market, i.e. a larger number of firms in the market, reduces firms' output. Interestingly, however, firms produce more in the 64K generation when the 16K generation becomes more competitive.

Our estimation results also reconfirm the significant intragenerational learning effects. The coefficients of the own cumulative outputs in the current generation is positive and highly significant. This illustrates the fact that firms raise production in order to gain more production experience which reduces own marginal costs via learning effects. In terms of intragenerational spillovers, other firms' cumulative output within the 64K generation (x_{-i}^{64K}) has an impact which is not significantly different from zero. This result seems to be surprising as Table 2 shows that intragenerational spillovers are prevalent. The estimation results of the output policy can be explained by the existence of two opposing effects: a direct effect having an impact on a firm's production via spillovers, and an indirect effect, as a firm produces less when other

¹⁸We also estimated the output equations using an instrumental variable estimator to correct for a potential endogeneity of the lagged own cumulative output. As instruments we use the twice and three times lagged variables of own cumulative output. The differences between the estimated coefficients from the instrumental variable and the ordinary least squares estimations are negligible. In addition, the Hausman test fails to reject the consistency of the ordinary least squares estimates. Moreover, the Hausman test statistics of 97.76 reject the random effects specifications.

firms' marginal costs decline. From our estimation results, both effects cancel each other out.

The intergenerational own learning effects (x_i^{16K}) and spillover (x_{-i}^{16K}) are insignificant. This result is consistent with the marginal cost estimates. Finally, the estimate of the coefficient of $\hat{\nu}_i$ is negative and significant as expected. A positive shock which rises marginal cost will reduce output.

Entry and exit policy We estimate the probabilities of entry, equation (9), and exit, equation (10), using probit models. The results are shown in Table 4. The first column reports the results of the entry policy. The results show that firms are more likely to enter when more firms have already entered the $64K$ generation in the past. The significant and positive estimate of cumulative output in the $16K$ generation emphasizes the importance of experience from earlier generations through intergenerational own learning effects. The estimated positive and significant time trend indicates that at identical industry states but different time periods, firms choose to enter in a later period in order to save on entry costs. This result provides a first insight that entry costs are decreasing over time.

Column 2 in Table 4 show the results for the exit policy.¹⁹ Firms are more likely to exit the $64K$ generation if they missed out on gaining sufficient production experience, and if many other firms entered the market. Moreover, firms are less likely to exit if other firms' cumulative outputs are larger.

5.2 Second stage results

Using our estimates, we apply forward simulations to calculate the discounted values (before entry costs) at different states given the optimal policy functions. We compare these discounted values to those generated from distorting the entry policies as described in Section 4.2. We retrieve the slope of the entry cost for the $64K$ generation using a simulated minimum distance estimator as shown in equation (13). The resulting entry costs are illustrated in Figure 7.²⁰ Our results indicate that entry costs decline by 342 million U.S. dollars during the first 30 periods. More than 90% of the reduction of entry cost occurs during the first 10 periods. In the first 10 periods, entry costs rapidly decline with an average of 31.8 million U.S. dollars per period.

¹⁹We do not include firm fixed effect in the exit policy since not all firms have exited the $64K$ generation at the end of data period

²⁰We normalized the entry costs presented in the figure by setting the entry cost of the 30^{th} period to 0.

The entry costs decline less steeply after the 11th period, which is the time when excessive entry begins (see Figure 4). Between the 11th and the 30th period, entry costs decline by about 1.2 million U.S. dollars per period on average. To summarize, our entry cost estimates support the fact that early entrants incur significantly higher entry costs.

5.3 Counterfactual experiment

Based on our structural estimates, we conduct a policy experiment that evaluates the impact of different lengths of entry regulation in the 64K generation on market structure and quantifies the impact on total welfare. We formulate the entry regulation as follows: The regulator grants a monopoly right, for example a patent, to the first firm entering the 64K SRAM chip generation. The right expires after a certain number of periods, which is specified by the regulator.²¹ The holder of the monopoly right receives the exclusive right to sell 64K chips for the time the regulation is being enforced. Once the entry regulation ends, potential entrants decide if they enter the chip generation, pay the corresponding entry costs and decide how much to produce.

Evaluating the impact of entry regulation on total welfare is not a straightforward exercise, as a longer regulation delays entry and also reduces the entry costs for firms. A lower entry cost has a significant impact on firms' discounted values, their entry decisions, as well as their quantity choices, prices and market structure. Firms are assumed to choose the same equilibrium strategies. We simulate output, prices, entry and exit depending on the different lengths of entry regulation enforced by the regulator. We also account for learning effects to evaluate the entire impact on consumer and producer welfare.

The results are illustrated in Figures 8 and 9. Note that all changes in total welfare are evaluated relative to the free entry case. Figure 8 shows the impact of the regulation on industry consumer and producer surplus. A longer entry regulation increases market power for the first entrant, which causes prices to increase, output to fall, and consumer surplus to decline. The change in producer surplus is always positive comparing to the case of free entry and increases as the length of regulation increases. The conclusion of welfare improving by entry regulation depends on the comparison of the increases in producer surplus and decreases in consumer surplus. For 3 and 4 periods of regulation, the total surpluses are larger than that under free entry. For 5 to 14 periods of regulation, the changes in total surplus are negative and continue to

²¹We assume that the original first entrant is being granted the patent. Hence, firms do not anticipate the regulation and other firms have no chance to outrun or leapfrog the original first entrant.

fall. This results reflect that the decreases in consumer surplus outweigh the increase in producer surplus. For 14 to 18 and 23 to 30 periods of regulation, the changes in total surplus rise due to the strong increase in the changes in producer surplus. While the losses in consumer surplus fall steadily and level off for longer entry regulation, the gains in producer surplus increase more rapidly for longer entry regulation. As a result, the changes in total surplus start to rise for longer regulation and become positive for more than 27 periods of entry regulations.

To better understand the change in producer surplus, we decompose the producer surplus into the monopolist's profits, other firms' profits and the savings of entry costs, see Figure 9. As the length of entry regulation increases, the protected monopolist's profits drastically increase relative to the free entry case for several reasons: First, protecting the first entrant for multiple periods preserves its monopoly position and enables the protected entrant to earn monopoly profits during these time periods. Under free entry other firms instantly enter the market. Hence, the change in market structure lowers the first entrant's profits. Since the entry regulation regime is compared to the free entry case, the monopolist's profits increase over time. Second, the monopolist produces the entire industry output and is able to further benefit from intragenerational own learning effects, which further lowers costs and increases profits. Comparing the profits gained by the monopolist to the profits lost by other firms, the monopolist's gain is lower than other firms' losses for shorter entry regulations (less than 15 periods) but greater than other firms' losses for longer entry regulations (more than 15 periods). With sufficiently long entry regulations, the protected monopolist can accumulate sufficient experience in early periods and accelerate the learning paces which enables other firms to benefit from spillovers. Consequently, the industry profits become positive relative to the free entry case for more than 15 periods of entry regulations.

The figure also show that entry cost savings increase with the number of monopoly periods, which is due to the fact that longer entry regulations delay successive entry. As the lengths of entry regulation increase, the increasing rate of entry cost savings depends on the decreasing rate of entry cost and the distribution of entry under free entry. When the entry cost declines faster and the entries of more firms have been delayed, the savings of entry cost are larger. From Figure 9, the savings of entry cost rise more rapidly for less than 10 periods of entry regulation because the entry cost decline drastically during the first 10 periods (see Figure 7). For more than 10 periods of entry regulation, the savings of entry cost increases relatively slowly

in general because of much lower decreasing rates of entry cost (see Figure 7). However, for 14 to 17 and 25 to 30 periods of entry regulations, the savings of entry cost rise more rapidly since the entries of a larger number of firms have been delayed by the regulation (see Figure 4). The entry cost savings are always larger than the change in industry profits (the combination of the monopolist's profits and other firms' profits) which makes the producer surplus to be positive relative to free entry case. For sufficiently long entry regulations (more than 25 periods), both the change in the monopolist's profit and savings of entry cost increase drastically and consequently lead to more rapid increases in producer surplus.

To summarize the main results of our policy experiment: for shorter entry regulations, there are losses in total welfare because of the losses in industry profit and small savings of entry costs. When the length of entry regulation lasts for sufficiently long periods, the large savings in entry costs plus the gain in monopolist profits due to learning effect rise producer surplus. Consequently, the gains in producer surplus outweigh the losses in consumer surplus and total welfare improves. Hence, free entry drives too many firms entering too early and therefore causes excessive entry.

6 Summary and concluding remarks

The purpose of this paper is to empirically investigate the impact of different entry regulations on consumer and producer surplus. We estimate a dynamic structural model in which firms choose the optimal time to enter a market, and make production and exit decisions. Our estimation results show that entry costs drastically decline over time. In fact, we find that entry costs decline by more than 90% in the first 10 quarters. Our estimation results also provide evidence that firms account for learning effects when determining their optimal production. Early entrants incur higher entry costs, but are able to take advantage of learning effects early in the life cycle, which lowers their future costs and returns higher profits. Late entrants especially benefit from decreasing entry costs over time as well as from other firms' experience via spillovers. Our policy experiments show that declining entry costs over time play a dominant role in evaluating the impact of entry regulation on total welfare. We find that consumer surplus is always declining for different lengths of entry regulations due to monopoly pricing. On the other hand, entry regulations increase producer surplus compared to free entry. Our results also show that different lengths of entry regulation increase total surplus in the industry if the duration of the regulation

is sufficiently short or long. If the regulation lasts shorter periods, total welfare decreases relative to the free entry case as the savings in entry costs is not large enough to compensate for the decline in consumer welfare and the loss in industry profit. If the regulation lasts sufficiently long, total surplus increases as the entry cost savings that result from avoiding excessive entry in the industry and the gain in the monopolist's profit compensate the losses in consumer welfare. Free entry results in excessive entry since firms under free entry only care about their own profits and not about the overall industry profits. Hence, entry regulation diminishes the private incentives of entry, and optimizes social incentives by avoiding excessive entry and firms paying a socially undesirable amount on entry costs. Even though our policy experiment suggests that there is excessive entry in the industry, and entry regulation would increase total surplus, we would like to remind that consumer surplus diminishes in our policy experiments. Therefore, we are cautious at this point to recommend entry regulations, and further research is warranted in this area.

On a final note, it is worth emphasizing that declining entry costs over time are significant also in many other markets, such as pharmaceuticals, the automotive and electronics industry, and many others. Therefore, it would be interesting to explore in future research if the main finding, i.e., sufficiently long entry regulations positively impact total welfare through avoiding excessive entry rates, also applies to other markets.

References

- Akerberg, D. and G. Gowrisankaran, 2006, "Quantifying Equilibrium Network Externalities in the ACH Banking Industry," *Rand Journal of Economics*, 37, pp. 738-761.
- Anderson, S.P., A. DePalma and Y. Nesterov, 1995, "Oligopolistic Competition and the Optimal Provision of Products," *Econometrica*, Vol. 63, pp. 1281-1302.
- Arrow, K., 1962, "Economic Welfare and the Allocation of Resources for Invention" in *National Bureau of Economic Research Conference Volume, The Rate and Direction of Inventive Activity*, (Princeton, NJ: Princeton University Press, 1962), pp. 609-625.
- Bajari, P., L. Benkard and J. Levin, 2007, "Estimating Dynamic Models of Imperfect Competition," *Econometrica*, Vol. 75, No. 5, pp. 1331-1370.
- Berry, S.T., 1992, "Estimation of a Model of Entry in the Airline Industry," *Econometrica*, Vol. 30, pp. 889-917.
- Berry, S.T. and J. Waldfogel, 1999, "Free Entry and Social Inefficiency in Radio Broadcasting" *The Rand Journal of Economics*, Vol. 30, pp. 397-420.
- Berry, S.T. and J. Waldfogel, 2001, "Do Mergers Increase Product Variety: Evidence from Radio Broadcasting," *Quarterly Journal of Economics*, Vol. 116, pp. 969-1007.
- Bresnahan, T.F. and P.C. Reiss, 1990, "Entry in Monopoly Markets," *Review of Economic Studies*, Vol. 57, No. 4, pp. 531-553.
- Bresnahan, T.F. and P.C. Reiss, 1991, "Entry and Competition in Concentrated Markets," *Journal of Political Economy*, Vol. 99, No. 5, pp. 977-1009.
- Campbell, J.R., and H.A. Hopenhayn, 2005, "Market Size Matters," *The Journal of Industrial Economics*, Vol. 53, No. 1, pp. 1-25.
- Carlton, D.W., 1983, "The Location and Employment Choices of New Firms: An Econometric Model with Discrete and Continuous Endogenous Variables," *The Review of Economics and Statistics*, Vol. 65, No. 3, pp. 440-449.
- Chamberlin, E., 1933, "The Theory of Monopolistic Competition," Cambridge, Mass.: Harvard University Press.

- Chaudhuri, S., P.K. Goldberg and P. Jia, 2006, "Estimating the Effects of Global Patent Protection in Pharmaceuticals: A Case Study of Quinolones in India, " *The American Economic Review*, Vol. 96, No. 5, pp. 1477-1514.
- Davis, P., 2006, "Spatial Competition in Retail Markets: Movie Theaters, " *The Rand Journal of Economics*, Vol. 37, No. 4, pp. 964-982.
- Dixit, A.K., and J.E. Stiglitz, 1977, "Monopolistic competition and Optimal Product Diversity, " *The American Economic Review*, Vol. 67, pp. 297-308.
- Dunne, T., M.J. Roberts and L. Samuelson, 1988, "Patterns of Firm Entry and Exit in U.S. Manufacturing Industries, " *The Rand Journal of Economics*, Vol. 19, pp. 495-515.
- Einav, L., 2010, "Not All Rivals Look Alike: Estimating an Equilibrium Model of The Release Date Timing Game," *Economic Inquiry*, 48(2), pp. 369-390.
- Ericson, R. and A. Pakes, 1995, "Markov-Perfect Industry Dynamics: A Framework for Empirical Work," *Review of Economic Studies*, 62, pp. 53-82.
- Fink, C., 2000, "How Stronger Patent Protection in India Might Affect the Behavior of Transnational Pharmaceutical Industries," *The World Bank, Policy Research Working Paper* 2352.
- Fudenberg, D. and J. Tirole, 1983, "Learning by Doing and Market Performance," *Bell Journal of Economics*, Vol. 14, pp. 522-530.
- Genesove, D. 1999, "The Adoption of Offset Presses in the Daily Newspaper Industry in the United States," *NBER Working Paper* No. 7076.
- Geroski, P. 1989, "Entry, Innovation and Productivity Growth," *The Review of Economics and Statistics*, Vol. 71, pp. 572-578.
- Gilbert, R.J. and C. Shapiro, 1990, "Optimal Patent Length and Breath," *Rand Journal of Economics*, 21, pp. 106-112.
- Gowrisankaran, G. and J. Stavins, 2004, "Network Externalities and Technology Adoption: Lessons from Electronic Payments," *Rand Journal of Economics*, 35, pp. 260-276.

- Hsieh, C.-T. and E. Moretti, 2003, "Can Free Entry Be Inefficient? Fixed Commissions And Social Waste In The Real Estate Industry, " *Journal of Political Economy*, v111(5,Oct), pp. 1076-1122.
- Hughes, J.W., M.J. Moore and E.A. Snyder, 2002, "Napsterizing Pharmaceuticals: Access, Innovation, and Consumer Welfare," *National Bureau of Economic Research Working Paper* 9229.
- Irwin, D.A. and P.J. Klenow, 1994, "Learning-by-Doing Spillovers in the Semiconductor Industry," *Journal of Political Economy*, Vol. 102(6), pp. 1200-1227.
- Klepper, S., 2002, "Firm Survival and the Evolution of Oligopoly," *Rand Journal of Economics*, Vol. 33, pp. 37-61.
- Klette T.J. and S. Kortum, 2004, "Innovating Firms and Aggregate Innovation, " *Journal of Political Economy*, Vol. 112, No. 5, pp. 986-1018.
- Lerner, J., 2002, "150 Years of Patent Protection," *American Economic Review*, American Economic Association, Vol. 92, pp. 221-225.
- Levin, R.C., A.K. Klevorick, R.R. Nelson and S.G. Winter, 1987, "Appropriating the Returns from Industrial Research and Development," *Brookings Papers on Economic Activity*, 3, pp. 783-820.
- Mankiw, N.G. and Whinston, M.D., 1986, "Free Entry and Social Inefficiency," *Rand Journal of Economics*, Vol. 17, pp. 48-58.
- Maskus, K.E. and D.E. Konan, 1994, "Trade-Related Intellectual Property Rights: Issues and Exploratory Results," In *Analytical and Negotiating Issues in the Global Trading System*, ed. Alan V. Deardorff and Robert M. Stern, 401-46. Ann Arbor, MI: University of Michigan Press.
- Mazzeo, M., 2002, "Product Choice and Oligopoly Market Structure," *The Rand Journal of Economics*, Vol. 33, No. 2, pp. 221-242.
- Moser, P., 2005, "How Do Patent Laws Influence Innovation? Evidence from Nineteenth-Century World Fairs," *The American Economic Review*, Vol. 95 (4), pp. 1215-1236.

- Nordhaus, W.D., 1969, "Invention, Growth, and Welfare," (Cambridge: MA: MIT Press.)
- Perry, M.K., 1984, "Scale Economies, Imperfect Competition, and Public Policy," *Journal of Industrial Economics*, Vol. 32, pp. 313-330.
- Riordan, M.H., 1992, "Regulation and Preemptive Technology Adoption," *Rand Journal of Economics*, Vol. 23, No. 3, pp. 334-349
- Schmidt-Dengler, 2006, "The Timing of New Technology Adoption: The Case of MRI," manuscript.
- Seim, K., 2006, "An Empirical Model of Firm Entry with Endogenous Product-Type Choices,"
" *The RAND Journal of Economics*, Vol. 37, No. 3, pp. 619-640
- Siebert, R.B. and C. Zulehner, 2013, "The Impact of Innovation and Market Demand on Market Structure," manuscript.
- Spence, A.M., 1976a, "Product Selection, Fixed Costs, and Monopolistic Competition," *Review of Economic Studies*, Vol. 43, pp. 217-236.
- Spence, A.M., 1976b, "Product Differentiation and Welfare, " *American Economic Review*, Vol. 66, pp. 407-414.
- Subramanian, A., 1995, "Putting Some Numbers on The TRIPS Pharmaceutical Debate," *International Journal of Technology Management*, Vol 10, pp. 252-268.
- Sutton, J., 1991, "Sunk Costs and Market Structure," MIT Press.
- Sutton, J., 2001, "Technology and Market Structure: Theory and History," MIT Press.
- Sweeting, A., 2006, "Coordination, Differentiation and the Timing of Radio Commercials," *Journal of Economics and Management Strategy*, Vol. 15, No. 4, pp. 909-942.
- Sweeting, A., 2009, "The Strategic Timing of Radio Commercials: An Empirical Analysis Using Multiple Equilibria," *Rand Journal of Economics*, Vol. 40, No. 4, pp.710-742.
- Toivanen, O. and M. Waterson, 2000, "Empirical Research on Discrete Choice Game Theory Models of Entry: An Illustration," *European Economic Review*, Vol. 44(4-6), pp.985-992.
- Toivanen, O. and M. Waterson, 2005, "Market Structure and Entry:Where's the Beef?," *Rand Journal of Economics*, Vol. 36, No.3, pp.680-699.

- von Weizsacker, C.C., 1980, "A Welfare Analysis of Barriers to Entry," *Bell Journal of Economics*, Vol. 11, pp. 399-420.
- Wattal, J., 1999, "Introducing Product Patents in the Indian Pharmaceutical Sector: Implications for Prices and Welfare" *World Competition, Law and Economics Review*, Vol. 20, No. 2, pp. 5-21.
- Wattal, J., 2000, "Pharmaceutical Patents, Prices and Welfare Losses: Policy Options for India under the WTO TRIPS Agreement" *The World Economy*, Vol. 23, No.5, pp. 733-752.
- Wright, B.D., 1983, "The Economics of Invention Incentives: Patents, Prizes, and Research Contracts, " *The American Economic Review*, Vol. 73, pp. 691-707.

A Appendix: Tables

Table 1: Demand Estimation Results

Variable	OLS	IV	OLS	IV
Constant	16.706*** (19.88)	18.728*** (21.77)	16.903*** (20.77)	17.406*** (24.04)
$\ln(P)$	-2.745*** (-25.77)	-3.187*** (-22.15)		
$\ln(P^{16K})$			-2.717*** (-13.47)	-3.067*** (-18.24)
$\ln(P^{64K})$			-2.579*** (-28.95)	-2.751*** (-29.31)
$\ln(P^{256K})$			-2.804*** (-20.57)	-2.927*** (-23.62)
$\ln(P^{1M})$			-3.006*** (-21.53)	-3.185*** (-22.71)
Dummy 64K	2.041*** (16.20)	2.194*** (17.32)	1.846*** (8.45)	1.661*** (8.44)
Dummy 256K	3.768*** (20.15)	4.149*** (22.52)	3.951*** (14.65)	3.742*** (15.44)
Dummy 1M	5.201*** (16.15)	5.942*** (18.37)	5.889*** (13.43)	5.803*** (13.80)
$\ln(P^S)$	0.345* (1.74)	0.336** (2.02)	0.185 (1.01)	0.436** (2.43)
$\ln(GDP)$	10.095 (0.79)	7.984 (0.60)	11.546 (0.95)	8.274 (0.71)
$\ln(P^{Dram})$	1.132*** (10.48)	1.223*** (10.86)	1.267*** (10.31)	1.217*** (9.92)
Time	0.013* (1.74)	-0.00003 (-0.00)	0.023** (2.54)	0.014 (1.53)
Number of observations	314	310	314	310
Adjusted R-squared	0.858	0.817	0.868	0.844

The dependent variable is the logarithm of industry output ($\ln(q_t^k)$). As instruments, we use the price of silicon, cumulative total output in the current generation, and cumulative total output from the previous generation. The t-statistics are shown in parentheses below the parameter estimates, and *** (**, *) denotes a 99% (95%, 90%) level of significance.

Table 2: Marginal Cost Estimation Results

Variable	(1)	(2)
$\ln(P^{stl})$ (ρ_1)	1.008*** (67.83)	0.729*** (36.28)
Intragenerational own learning (ρ_2)	-0.344*** (-61.00)	-0.260*** (-51.76)
Spillover learning (ρ_3)	0.317*** (6.14)	0.308*** (5.19)
Intergenerational own learning (ρ_4)	0.010*** (4.63)	0.0009** (2.46)
Dummy 64K	0.239*** (21.02)	0.293*** (25.96)
Dummy 256K	0.639*** (52.10)	0.731*** (57.62)
Dummy 1M	1.008*** (67.23)	1.204*** (69.54)
Time		-0.013*** (-20.46)
Firm fixed effect	Yes	Yes
Number of observations	5,835	5,835
Adjusted R-squared	0.907	0.913

Table 2 shows nonlinear least square results of the marginal cost function, as shown in equation (4). The dependent variable is the logarithm of the marginal costs. The t-statistics are shown in parentheses below the parameter estimates, and *** (**, *) denotes a 99% (95%, 90%) level of significance.

Table 3: Output Policy Estimation Results

Variable	64K
Constant	-5.971 (-1.04)
GDP	0.00000038*** (3.25)
$\ln(P^{DRAM})$	0.200** (1.98)
$\ln(P^{sil})$	-0.391 (-1.48)
$\ln(n^{16K})$	0.430** (2.12)
$\ln(n^{64K})$	-0.296 (-1.30)
$\ln(x_i^{16K})$	-0.044 (-1.17)
$\ln(x_i^{64K})$	0.492*** (11.34)
$\ln(x_{-i}^{16K})$	0.643 (1.14)
$\ln(x_{-i}^{64K})$	0.013 (0.07)
$\hat{\nu}_i$	-0.319** (-2.57)
Time	-0.094*** (-6.64)
Firm fixed effect	Yes
Number of observations	1,693
Adjusted R-squared	0.782

The dependent variable is the logarithms of firm level output in the 64K generation. The t-statistics are shown in parentheses, and *** (**, *) denotes a 99% (95%, 90%) level of significance.

Table 4: Entry and Exit Policy Estimation Results

Variable	Entry 64K	Exit 64K
Constant	-4.680*** (-5.58)	-2.658*** (-3.62)
$\ln(n^{64K})$	0.170*** (2.71)	1.104* (1.91)
$\ln(x_i^{16K})$	0.046** (2.31)	
$\ln(x_i^{64K})$		-0.139*** (-4.43)
$\ln(x_{-i}^{64K})$	0.051 (1.63)	-0.395* (-1.71)
Time	0.054*** (7.10)	0.051*** (2.70)
Firm fixed effect	Yes	No
Number of observations	3,195	1,701
Pseudo R-squared	0.592	0.058

Table 4 shows the results for the entry and exit policies, as shown in equations (9) and (10). The entry and exit policies are estimated using probit models. The dependent variable in the entry model takes on a value of one when a firm chose to enter the 64K generation and zeros before entry occurred. In the exit model, the dependent variable takes on a value of one if a firm exits the generation and zero between the firm's entry and exit decisions. Firm fixed effects are not included in the exit models since not all firms have exited in the data. The z-statistics are shown in parentheses, and *** (**, *) denotes a 99% (95%, 90%) level of significance.

B Appendix: Figures

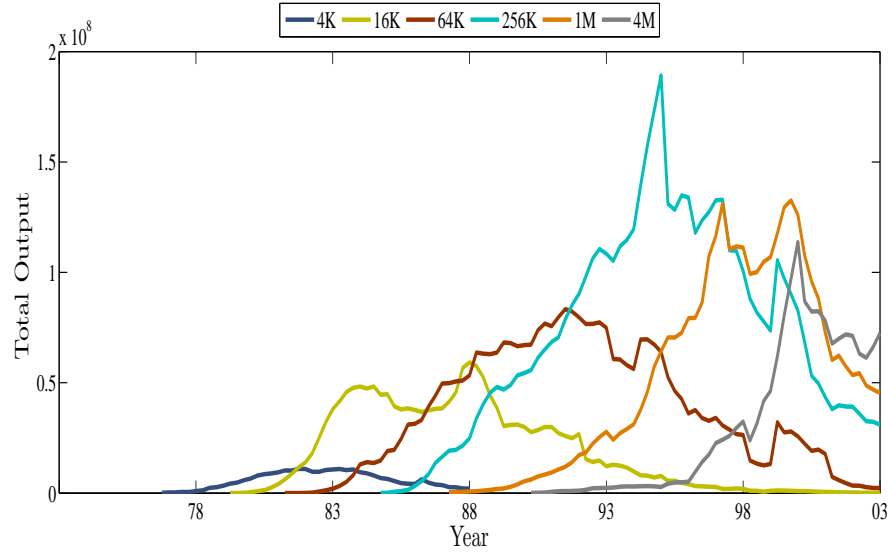


Figure 1: Industry units shipped, 1974-2003
Source: Gartner Inc.

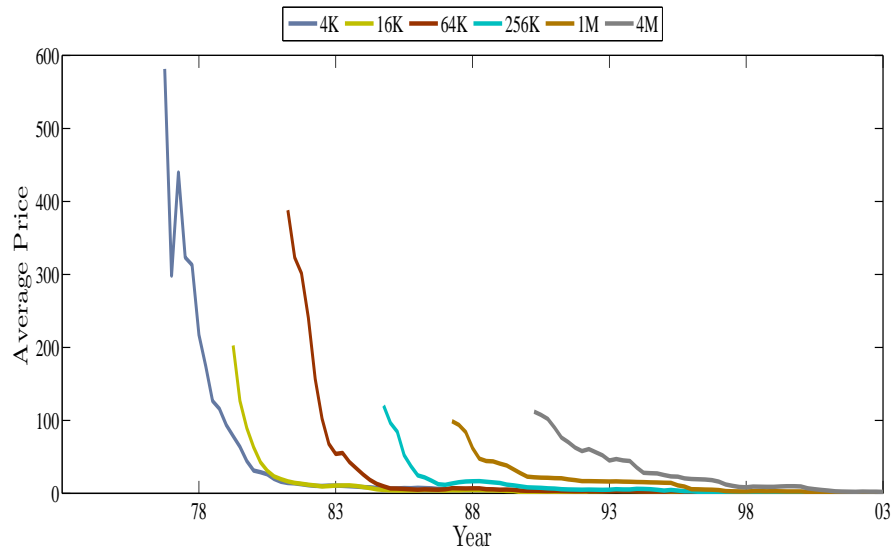


Figure 2: Average SRAM prices, 1974-2003
Source: Gartner Inc.

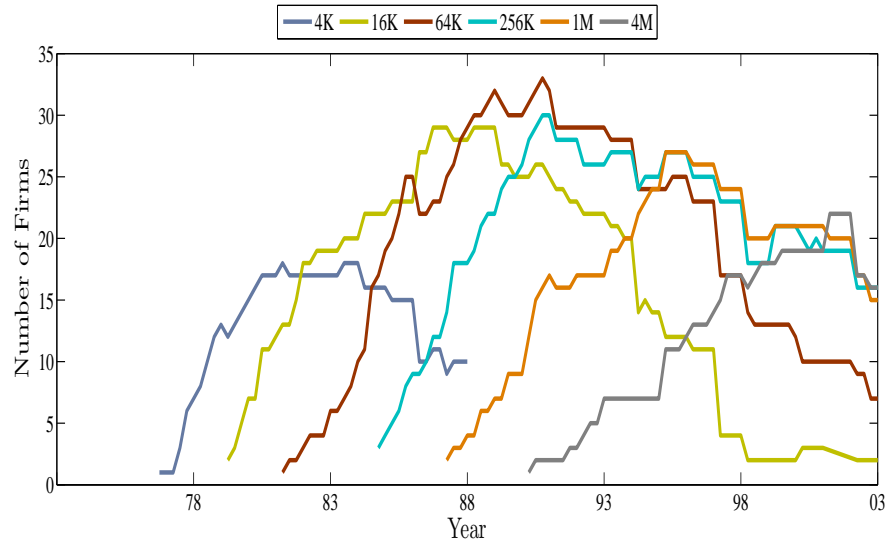


Figure 3: Number of Firms
Source: Gartner Inc.

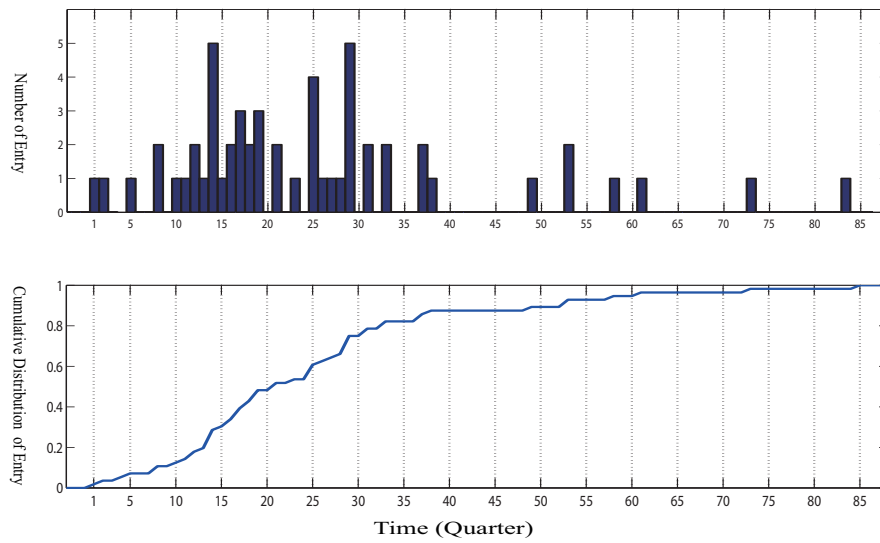


Figure 4: Distribution of Entry 64K
Source: Gartner Inc. The first entry occurs in the first quarter of 1982.

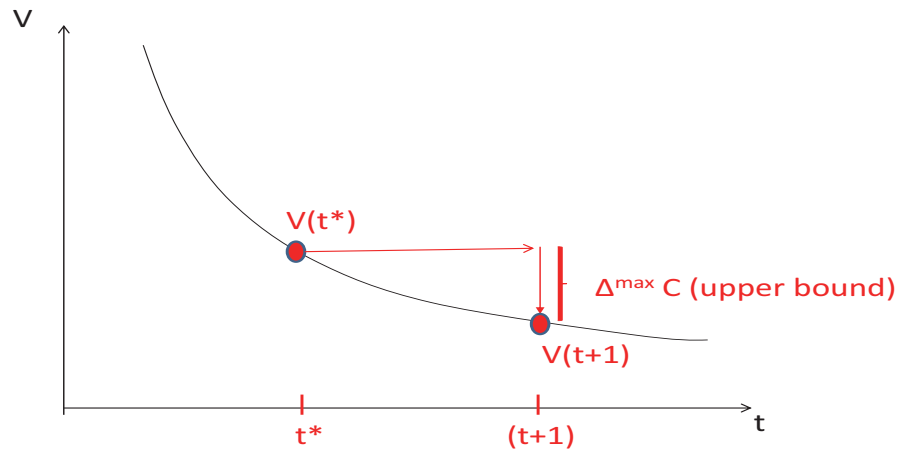


Figure 5: Identification of Entry Cost (Upper Bound)

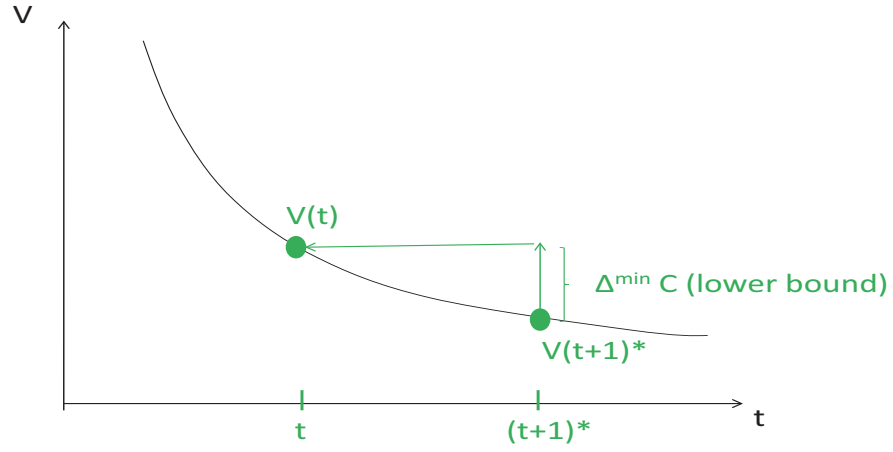


Figure 6: Identification of Entry Cost (Lower Bound)

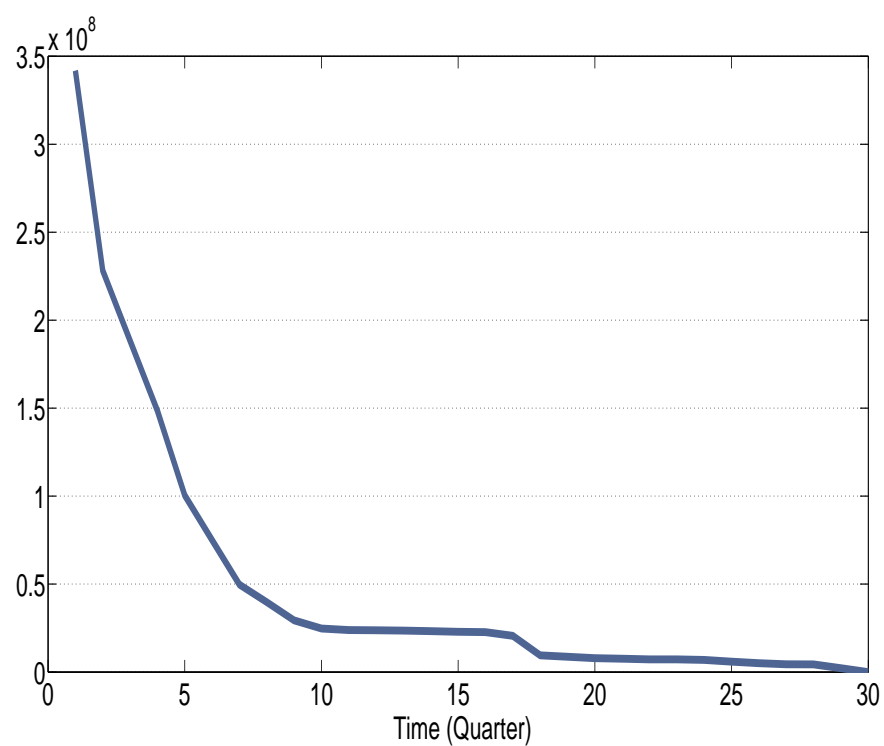


Figure 7: Slope of Entry Costs for the 64K generation

All values are the discounted values at the timing of the first entry in US dollars.

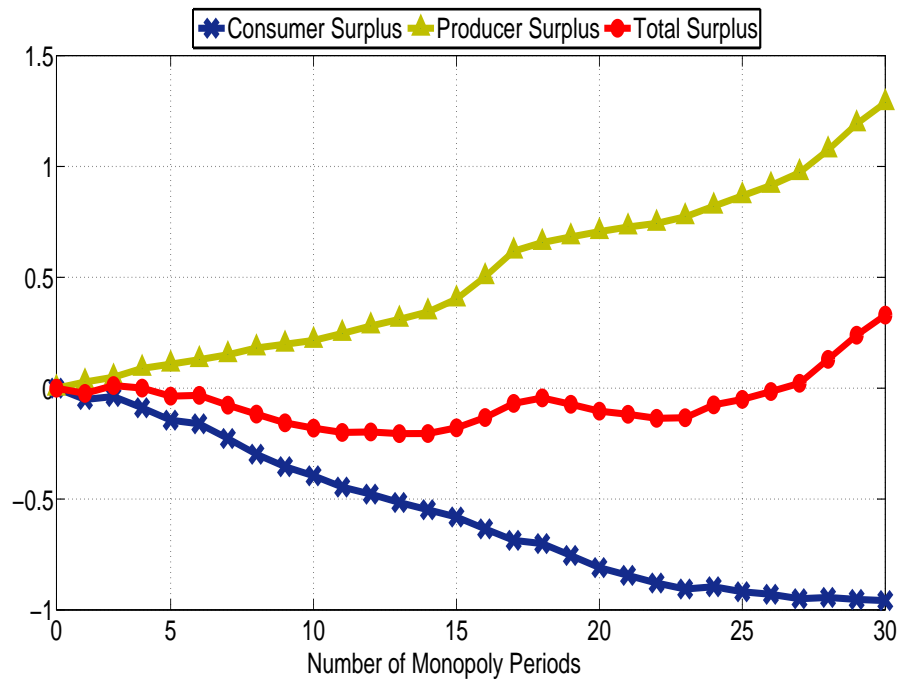


Figure 8: Change in Welfare

All values are the discounted values at the timing of the first entry in US dollars.

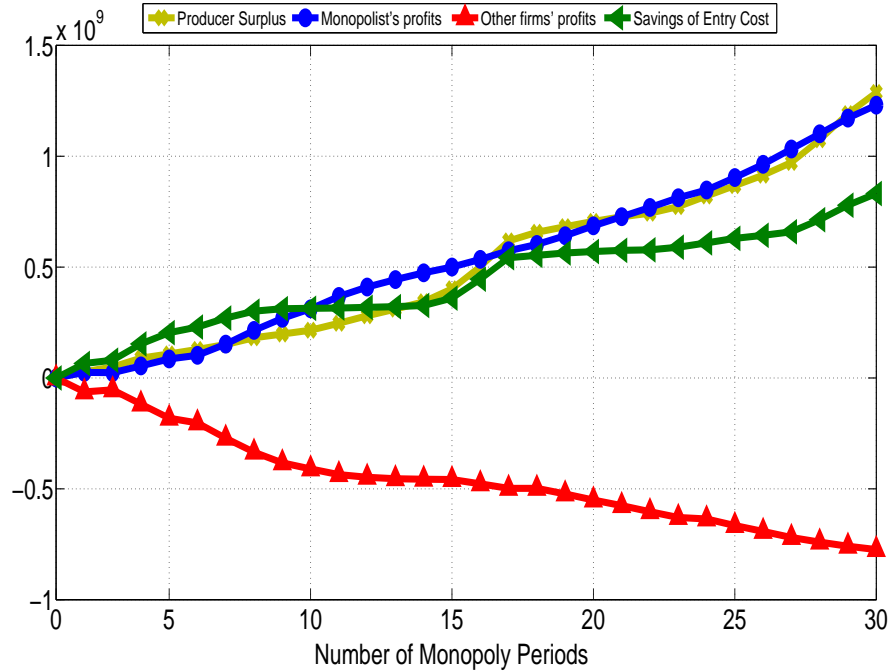


Figure 9: Change in Producer Surplus

All values are the discounted values at the timing of the first entry in US dollars.