The Impact of Innovation and Market Demand on Market Structure

Ralph Siebert* Christine Zulehner†

July, 2008

Abstract

We study why the number of firms in the dynamic random access memory semiconductor industry is drastically increasing throughout early memory generations, whereas sharply decreasing for recent generations. This fact is even more surprising since market demand steadily increased over time whereas the pace of innovation continously decreased over time. We estimate a fully dynamic oligopoly model accounting for entry, exit, and learning by doing. We account for serially correlated observed and unobserved state variables and apply the recent two-step estimator by Bajari, Benkard and Levin (2007) using quarterly firm-level data from 1974 to 2004. We find that the interdependence between the increase in market demand and rapidly increasing sunks costs associated with a higher pace of innovation nicely explains the change in market structure. We also confirm that accounting for serially correlated unobservables is a crucial aspect to correct for. In our case, the instrumental variable estimator with a lagged dependent variable performs best when accounting for a serially correlated unobserved state variable.

JEL: C1, L1, L6, O3.

Keywords: Dynamic Random Access Memory, Dynamic Model, Discrete Choices, Entry, Exit, Innovation, Learning by Doing, Semiconductors, Sunk Costs.

Acknowledgments: We would like to thank Jason Abrevaya, Daniel Ackerberg, Victor Aguirregabiria, Pascal Courty, David Genesove, Gautam Gowrisankaran, Saul Lach, Michael Mazzeo, Amil Petrin, Stephen Ryan and Philipp Schmidt-Dengler for helpful discussions and valuable suggestions. We gratefully acknowledge financial support from the Jubiläumsfonds of Oesterreichische Nationalbank (grant 11038). Christine Zulehner also gratefully acknowledges financial support from the Austrian Science Foundation (grant J2481-N12) and thanks the Economics Department at Purdue University for its hospitality where the paper was partly written. All errors are our own.

*Purdue University, Krannert School of Management, Economics Department, 403 West State Street, IN 47906-2056, USA, Email: rsiebert@purdue.edu.

†University of Vienna, Department of Economics, BWZ-Bruennerstrasse 72, A-1210 Vienna, Email: christine.zulehner@univie.ac.at.

1 Introduction

Many policy debates target on evaluating the competitiveness in new product industries, such as the the semiconductor, or the dynamic random access memory (DRAM) industry.[1] The DRAM industry is considered a new product industry which experiences a high pace of innovation that leads to bringing new product generations on the market, and high entry and exit rates.

Entry and exit may cause considerable consequences on market performance. Entry may increase the efficiency and may also accelerate the competitive pressure and pace of innovation in the market having a positive effect on efficiency and market performance. A higher innovation rate may also drive firms out of the market fewer survivors will lead to an increased market concentration, a lower degree of competition and may reduce the innovation rate. On the other hand, exit may also lead to efficiency gains, as the output formerly produced by the extinguished firms will be reallocated to the surviving and more efficient firms, see also Salant and Shaffer (1999). Hence, the pace of innovation and the turnover of firms is an important topic to explore from a policy point of view.

The number of firms increased from 15 firms in the 4K DRAM generation in 1978 to about 24 firms in the 4MB generation in the mid 1990's. Afterwards, the number of firms sharply declined to 15 firms in the 128 MB in 2001, and declined even further to 11 firms for the 256 MB generation in 2004.[2] The question is why is the DRAM industry characterized by this inverse U-shape in number of firms over different generations?

The U-shape in the number of firms is especially surprising because of two reasons. First, it is well known that the DRAM industry is characterized by an increasing pace of innovation over time. The invention of a new technology is a necessary condition for independently introducing a new chip generation and, hence, part of the entry or sunk cost. New electronic products imposed higher requirements towards the capacity of DRAM chips and new DRAM technologies became more complex over time, requiring higher R&D investments. For example, as shown in Table 1a, the number of patent applications in the DRAM industry increased from 701 in 1989 to 2,390 in 1997. The usage of simple patent counts may roughly indicate the increasing pace of innovation. However, it is controversial whether patent counts are good proxies for innovation and it is even less clear to what extent patent counts will approximate generation-specific R&D investments and entry or sunk costs. Very little is known about the entry or sunk cost of a new

[1] Dynamic Random Access Memories are components within the semiconductor industry. For more information on the industry, see the Industry Description.

[2] Note that the 128 MB and the 256 MB DRAM generations passed the maximum of industry production at the time periods that we are refering to. This would approximate the maximum number of firms in these generations, as there is not much more entry occuring afterwards.

generation, e.g. there a no generation-specific R&D investments reported. Using patent counts as a proxy for investment costs is controversial in iteself and would also be problematic since we would not know what lag strcture we had to assume.

Few sources report that the establishing costs of a plant and the evolution of the innovation cost for generations over time. As shown in Table 1b, a plant with a capacity of 30,000 chips per month rose from US-$ 1/2 billion in 1985 to US-$ 2.5 billion in 1999, and to about US-$ 5 billion in 2007. The problem with using those costs is similar to using production costs: they are rarely observed and may not be reliable for accounting reasons. Consequently, the first surprising fact is that the increasing pace of innovation may explain the recent declining number of firms because firms may not be able to generate enough profits in order to be able to reinvest those in the development of future technologies. However, it rather conflicts the fact that the number of firms increased in the early generations even though the R&D requirements increased.

The second reason why a U-shape in the number of firms is a surprising observation is given by the fact that many downstream industries permanently relied on DRAMs as inputs for electronic devices such as cell phones, computers, and video games, among many other devices. Whereas the ongoing growth in demand for DRAMs explains the increasing number of DRAMs firms for early generations very well, it is not as straightforward to find an explanation for the decreasing number of firms for more recent generations.

Recent literature emphasizes the interdependence between those factors, in order to evaluate the competitiveness in those markets. Various models have been proposed that focus on the interdependence between innovation, entry, and exit, see also Audretsch and Klepper (2000) and the literature cited therein for a nice overview in this area. Geroski (1996) distills a series of "stylized facts and results" from the empirical literature on entry. He concludes that entry is less a mechanism for keeping prices down through competitive pressure and more a mechanism for bringing about change associated with innovation. Moreover, it has been shown that exit rates are higher in more innovative industries. Katz (2007) also highlights the relevance to distinguish between short run and long run effects on market structure and therefore strengthens the importance of accounting for the interdependence of those aspects. Hence we are also interested in capturing and evaluating the interdependence of those effects on market structure.

We are interested in evaluating to what extent demand conditions, increasing sunk or entry costs and a higher pace of innovation drive entry and exit in the market. We find that growth in market demand and sunk costs increased over different generations. The demand increased by 3% from 1976 until late 90s and reducded to 1% afterwards. Our sunk cost estimates seem to be quite reliable as they are getting close to the reported establishment costs over different generations, see also Table 1b. The sunk costs were increasing over time. To summarize, our results show that the growth in demand dominates the growth in sunk entry cost up until the mid 1990's. Consequently, firms entered the DRAM market. After the mid 1990's, however, sunk costs dominated the growth in market demand. Since the losses from increasing sunk costs dominates the benefits from diminishing growth rates. Hence, the inverse U-shape in the industry can be explained by the interdependence between the diminishing growth in market demand and the increase in sunk costs.

Since we target on the evolution of one market over time, the analysis of entry and exit requires a dynamic model in which sunk costs play an important element.[3] One main difficulty in answering these questions is given by the fact that generation-specific R&D investments or entry costs are either rarely available, or not reliable. The question arises, how we can gain information on the change of sunk costs over time without having appropriate data? Even if we considered sunk costs to be exclusively determined by innovation, and even if we assumed sunk costs would perfectly be correlated with patent information, we would still face problems such as, how many years we had to lag the patent data, and how to appropriately transfers the patent informatoin into a monetary value.

The main contribution of our study is the estimation of sunk cost in a fully dynamic oligopoly model. We formulate a dynamic game in the tradition of Ericson and Pakes (1995) in which forward looking firms make entry, exit and production decisions and maximize their expected discounted sum of profits over the life cycle. We infer the entry or sunk costs from firms' equilibrium behavior. As our dataset consists of discrete as well as continous information, we estimate this model using the two step estimator by Bajari, Benkard, and Levin (2007), in which we implement an additional serially correlated unobserved state variable, e.g. firm-specific productivities.[4] In the first step, the policy functions, such as entry, exit and production will be estimated. In the second step we estimate the structural parameters such as the sunk costs.[5]

[3]Pioneering literature on entry and exit applied static approaches in order to evaluate the number of firms that different markets can hold depending on their market sizes and fixed costs, see e.g. Bresnahan and Reiss (1989, 1990a,b, 1991), Seim (2006), Asplund and Nocke (2006) and Waterson and Toivanen (2005).

[4]Note that we are interested in analyzing the competitive degree in the DRAM market and would like to estimate the entry and exit costs in a dynamic model allowing for observed and unobserved serially state variables. Examining responses to policy or environmental change, would be an interesting task as well, but goes beyond the scope of the paper.

[5]One strand of literature applied dynamic models that require to solve for Markov Perfect Nash Equilibria. Prominent examples that used the Pakes and McGuire (1994) algorithm are Benkard (2004), Gowrisankaran and

It is well known that learning by doing is an important phenomenon in the semiconductor industry, see e.g. Gruber (1996), Irwin and Klenow (1994), Siebert (2007), and Zulehner (2003). Firms learn through their experience in producing semiconductor wafers. Through repetitions and fine tuning of production processes, they are able to lower manufacturing costs. Past accumulated output is usually used to proxy firms experience.[6] It is assumed that firms slide down the same (industry) learning curve which illustrates efficiency effects that firms achieve in the long run through learning by doing. Hence, depending on every firms' past experience they are at different locations on the industry learning curve. Production enters firms' costs through experience and, therefore, becomes a state variable. We need to be aware of the fact that firms production has a contemporaneous impact of prices and profits, as well as an intertemporal impact on firms profits through their costs. Quantities and prices are not solely determined in static equilibrium, but also through intertemporal production plans. This makes it difficult to separately estimate firms' static profits from their continuation values. Since firm-level production is observed in our dataset it enters our model as an observed state variable.

However, we also want to account for firm specific productivities that may describe potential short run deviations from the common industry learning curve. Those firm-specific deviations could be occuring through shocks in the economy that firms may handle differently depending on their firm-specific productivities. Firm-specific productivities are stemming from the fact that firms have different capabilities to learn. They may occur e.g. through differences in managerial abilities, technological (absorptive) capacities, innovation ability, organizational structure, or strategic alliances.[7] Most importantly, they have an impact on firms costs, production decisions

Town (1997) and Pesendorfer (2003). Solving for Markov Perfect Nash Equilibria, however, are computational very complex as the continuation values need to be calculated for different parameter values using a nested fixed point algorithm. Therefore, the algorithm is very restricted with regard to the number of states and players.

Very recent studies focus on reducing the computational burden in dynamic games by estimating the continuation values and apply a two step algorithm. In the first stage a policy function is estimated and in the second stage the structural parameters will be recovered, see Aguirregabiria and Mira (2007), Bajari, Benkard and Levin (2007), Pakes, Ostrovsky and Berry (2006), Pesendorfer and Schmidt-Dengler (2006). For further discussion and an excellent description of the different methods, see also Ackerberg, Berry, Benkard and Pakes (2005). As of now there are only few studies that estimate a fully dynamic oligopoly model applying a two-step algorithm, exceptions are Bersteanu and Ellickson (2005), Collard-Wexler (2005), Gowrisankaran, Lucarelli, Schmidt-Dengler, and Town (2008), Macieira (2006), Ryan (2006), and Sweeting (2006). One common feature in those studies is that state variables are commonly observed by the players and the econometrician. However, very little is known about the implementation of additional unobserved state variables.

[6] note that for some industries such as the aircraft, or biotechnology industry the learning process might represent forgetting, see Bendkard and Argote. However, since the semiconductor industry is reprensented by cummulative innovation process and short life cycles, we abstract for forgetting.

[7] We could also interpret the LBD effects as long run or more persistent efficiency effects which reduce

and payoffs and therefore present a state variable. Since those firm-specific productivities are the outcome of firms investment in their capabilites, they should be allowed to vary over time. Firms that have been more productive in the past are more likely to be more productive today. Moreover, empirical studies frequently face the problem that firm-specific productivities are unobserved. Hence, we need to account for the firm specific productivity to be a serially correlated state variable.

As firm-specific productivities determine production decisions they may cause a simultaneity bias, as the contemporaneous correlation between the firm-level productivity and other regressors may cause a simultaneity bias. Different alternatives have been suggested to account for the simultaneity bias. Prominent studies in the production function literature account for time variant firm-specific productivity, or serially correlated unobserved state variables, by applying a proxy variable approach, see e.g. Olley and Pakes (1996), Levinsohn and Petrin (2003), Caves, Frazer and Ackerberg (2005) and Wooldridge (2005).[8] For example, Olley and Pakes (1996) assume that productivity is private information which is serially correlated and evolves over time according to an exogenous Markov Process. Under the assumption that investment is increasing in productivity, they use investment as a proxy for productivity. Their proxy controls for the part of the error term correlated with the endogenous regressors by absorbing any variation that is possibly related to the productivity term. They apply a semiparamentric estimation procedure using a polynomial series estimator in investment and capital. The problem with applying a proxy variable approach is that it is not easy to find appropriate or sufficient data in order to appropriately proxy for the firm-specific productivity.

Applying an instrumental variable approach is another alternative, e.g. searching for instruments that are highly correlated with the endogenous regressors, but not correlated with the productivity term. Appropriate instruments in the production function literature are firm-level factor prices or lagged inputs. The problem is that the former are rarely available and the latter are valid instruments only if the time series is long enough. To summarize, the decision whether to apply a proxy or an IV approach gets back to finding appropriate proxies for the omitted variable (productivity) or finding appropriate instruments for the endogenous regressors, respectively. Since we would face difficulties in finding a good proxy for firm-level productivity, we apply the instrumental variable approach. Having firm-level data over a sufficiently long time series reinforces the decision to follow the instrumental variable approach.

---

the marginal costs, whereas the serially correlated firm-level productivity differences.would represent short run efficiency effects that capture the firm-level variation around the long run effects.

[8] For a discussion of how to correct for serially correlated unobserved state variables, see also Bajari, Benkard, and Levin (2007), Pakes (1993), and Ackerberg, Berry, Benkard and Pakes (2005).

Very little is known about the treatment of serially correlated unobservables in dynamic games. The implementation of unobserved state variables makes identification a challenging task. Contrary to i.i.d. shocks, we explicitly need account for the fact that players and econometrician need to form beliefs over the distribution of their rivals' unobserved state variables. When solving this problem backwards, we get back to the initial condition problem. The regular procedure in solving this problem is to try different initial draws, or assuming a functional form for the unobserved characteristic. Some studies estimate dynamic models, but not games, allowing for autocorrelated errors, see e.g. Keane and Wolpin (1994), Stinebrickner (2000) and Bound, Stinebrickner, and Waidman (2005). Duflo, Hanna and Ryan (2007) estimate a structural dynamic labor supply model that allows for serial correlation in the opportunity cost of attending school. Their solution to the problem of serially correlated errors is to integrate out over the unknown distribution of the error term. To overcome the "initial condition" problem they make the assumption that agents receive an idiosyncratic draw from an unconditional error distribution. In our study, we can easily overcome the initial condition problem since every single DRAM generation starts from the same initial state, which is zero production.

We apply the two stage estimator by Bajari, Benkard and Levin (2007) as it allows us to incorporate continuous choices as well as discrete choices. Their estimator builds on the idea of Hotz, Miller, Saunders, and Smith (1994) to apply forward simulation for obtaining the continuation values. In the first step, we estimate the policy functions for entry, exit and production. The reduced-form policy functions describe what actions the firms take given the state. We assume that firms cost function is characterized by observed and unobserved serially correlated state variables, learning by doing and firm-level productivities, respectively. We assume that the time variant firm-specific productivity enters the firms cost function and follows a first order autoregressive process, such that it depends on the last period's productivity and an independent private shock every period. A positive autocorrelation indicates that more (less) productive firms remain more (less) productive for a while. Firms therefore need to form beliefs about their opponents productivity. The productivity enters firms production policies.

We apply four different estimators in the production policy in order to correct for the unobserved productivity. The first estimator treats the unobserved productivity as an error term that follows a first-order autoregressive process, also known as autocorrelation and apply a GLS estimator. We use instruments for learning by doing. We do not account for time invariant unobserved heterogeneity in this case, in order to avoid inconsistent estimats as the unobserved heterogeniety would be correlated with the past accumulated output. The second estimator directly controls for the serially correlated unobserved productivity by applying a lagged dependent variable model, or an AR(1) model. In order to correctly account for the dependence between the lagged and the current dependent variable, we also need to account for time invariant unobserved heterogeneity in order to avoid spurious correlation. We use instruments for the lagged dependent variable and learning by doing. The Problem with the IV estimator in levels is the initial condition problem. However, we can easily get around this problem as firms have the same initial condition at the beginning of every generation. The third estimator is a first difference estimator using instruments for the differenced learning variable. We apply the standard GMM estimator by Arellano and Bond (1991), which eliminates unobserved firm-specific effects by taking differences. The problem is that instruments are not strongy correlated as the series on production is highly persistent, so that lagged levels are only weakly correlated with first differences. We finally apply the estimator by Blundell and Bond (1998), which is similar to Arellano Bond (1991), but uses more efficient instruments.

We also estimate the policy functions without unobserved state variables. Our first stage estimates confirm that accounting for a correlated unobserved state variable gives signficantly different results in our application compared to not accounting for an unobserved state variable.

In a next step, we use forward simulations based on the optimal policy functions in order to generate more observations and calculate the continuation values given the optimal policies. Next, we distort the optimal policy function and establish alternative suboptimal policy functions from which the new suboptimal production paths are calculated.

In the second stage we use the structural parameters, such as generation-specific entry and exit costs, parameters from the cost functions and the distribution of private shocks. We use a simulated minimum distance estimator and look for those parameters that provide the best fit to the data that are generated by the optimal policies representing the equilibrium outcomes of profit-maximizing firms, compared to the data generated from suboptimal policies.

Fianlly, we calculate the sunk costs by calculating the expect discounted values at different states and comparing them to entry observations at those state. If entry occured at those states this means that sunk costs must be lower than the generated discounted profits generated at thsi stage.

The remainder of the paper is as follows. The next section provides an industry description providing some insight into the production technology as well as a first insight into the data. Section 3 introduces our dynamic oligopoly model and Section 4 presents the econometric model. In Section 5 we present the empirical results. We conclude in Section 6.

2 Industry Description and Data

Semiconductors are a key input for electronic goods, such as computers, consumer electronics, and communications equipment. The semiconductor industry is considered to be important as it has a significant impact on many other downstream industries with a high impact on growth. The semiconductor market consists of memory chips, micro components, and other components such as logic devices.

One element of the memory devices are DRAM chips which become more complex and higher requirements as tehy are used for games etc. More complex technology as electronic products more demandin in memory. DRAM are differentiated by memory capacit how much memeory can be stored on a wafer. Increasing demand on elelctr products increases the necessity to introdune new generations with higher memory capacity which increases the pace of innovation.

Even within one generation processes need to constatnyl be imporoved and the pace of innovation increases and the need to improve existing production processes increases over time. Figure 2 shows the industry shipments (in mio.) across different generations, i.e. the 4K till the 1GB generation from the years 1974 to 2004 on a quarterly basis. The figure illustrates that DRAM shipments follow a product life cycle, that lasts for approximately 5 years in the in 1980's to 3-4 years in the late 1990's. A higher pace of innovation in downstream industries and higher demand for more advanced chips is one reason why product cycles became shorter. Shorter life cycles put higher pressure on firms to recoup research and development cost within shorter time and generated profits may become too low for reinvesting into new product generations. Figure 2b illustrates the average, minimum and maximum industry units shipped in the DRAM industry over different generations. This figure clearly illustrate the upward trend in average shipments over time and emphasizes a continuing market growth.

The pace of innovation in this industry also increases due to improving existing production processes over time within a generation. DRAM chips are produced in batches on silicon wafers. The production process requires a complex sequence of photolithographic transfer of circuit patterns from photo masks onto the wafer and of etching processes. Regarding the manufacturing process within a generation, it has to be very precise in terms of temperature, dust, vibration levels and other determinants. It is of fundamental importance that this process occurs in clean rooms, as even tiny dust particles on the wafer surface interrupt the connecting pattern and thus the chip useless. The wafer, once processed, is cut and the single chips are then assembled.[9] Firms learn about production processes and devote an increasing share of

[9] More detailed descriptions of the production processes can be found in e.g. Gruber (1996), Irwin and Klenow (1994) and Flamm (1993).

their innovative activity to improving the production processes. The DRAM industry is characterized by extensive learning by doing effects, resulting from the fine-tuning of production processes. The yield rate, which is measured by the ratio of usable chips to the total number of chips on the wafer increases through learning over time. Against the background of learning by doing, firms' unit costs decline over time as production experience is accumulated through past output. Learning by doing is highest at the beginning of the life cycle and slow down over time. Given a decreasing rate of learning within a generation cost differences between firms become smaller throughout the life cycle which toughens price competition towards the end of the life cycle for a generation. Learning by doing is one reason why we need to account for a dynamic model as firms' current output will increase future experience which results in future cost savings. Given the existence of learning effects our study will account for the fact that firms follow a dynamic production strategy, such that firms' current production will reduce production costs in the future (see e.g. Dick, 1991; Fudenberg and Tirole, 1983; Majd and Pindyck, 1989; Spence, 1981; and Wright, 1936). Learning by doing is frequently used to explain the rapid price decline of the different generations (see Figure 1). Figure 1b which illustrates the average, minimum and maximum prices for the different DRAM generations. The price decline is an important aspect in the industry and even more so as it is transmitted to many other downstream industries having an impact on economic growth. The price decline looks quite similar for different generations. The price at which a new generation was introduced into the market, however, is quite different across generations, and does not follow an obviously comparable pattern. The learning by doing aspect is generation-specific, as production takes place in specific plants using specific production processes. Gruber (1992) also notes that learning enters the manufacturing process through the fine-tuning of generation-specific production processes. Irwin and Klenow (1994) confirmed the statement and found only low, sometimes even nonexisting intergenerational spillovers in the market. Consequently, plant-specific investments are driven by incurring generation-specific sunk costs.

We use firm level and industry level information on prices and quantities of different DRAM generations which are compiled by Gartner Inc. The data cover firm units shipped, industry units shipped, the average selling price, and the number of firms in the market from January 1974 to December 2004 on a quarterly basis. The dataset encompasses 14 generations, namely the 4K, 16K, 64K, 256K, 1MB, 2MB, 4MB, 8MB, 16MB, 64MB, 128MB, 256MB, 512MB, and 1GB generation.

Tables 2a and 2b shows the different producers for different generations. It is very interesting to note that the number of firms increases from 15 in the 4K DRAM generation to 22 firms in the 64K and 24 firms in the 4 MB and 23 firms in the16 MB DRAM generation. The number of firms then declines down to 15,11, 7, and 5 firms in the 128MB, 256MB, 512MB, and 1GB DRAM generation, respectively. See also Figure 3b and Figure 3c for the evolution of the number of firms over different generations. Note that the number of firms is not due to truncation problem or the fact that life cycles just started as for 256MB gerentaion hte peak of the life cycle has been reached. It is interesting to note that firms enter different DRAM generations successively, meaning that once a firm exits the market it will not reenter in any successive generation. Moreover, our dataset shows that firms enter new generations at most 2 years after the generation has been launched. Entry occurs (almost exclusively) at the beginning and exit mostly at the end of the life cycle. One explanation for finding few turbulences in the middle of a DRAM generations is consistent with learning effects, as late entrants may not be sufficiently efficient to compete with firms that are further down the cost function due to learning effects.

Our dataset also shows that the ranks for the top sellers between different generations is quite persistent. This observation might be an indication that firm-specific determinants such as productivity managerial talent or organizational structures are important features. This argument strengthens our idea to account for serially correlated unobserved state variables. Moreover, it is quite interesting to note that firms exiting and entering different generations occur at the lower ranks. It is the small firms who exit, and firms entering the DRAM market begin with a low production volume. These facts indicate that entry and exit decisions are long run decisions in which the entry and exit cost is related to firms' profits over the product life cycle.

Table 3 provides summary statistics of some of our variables that we will be using in our empirical analysis later on. In a first step, we will apply very preliminary regressions in order to investigate if learning effects are prevalent in our dataset. From the output data, we construct current industry output and past cumulative industry output. Current output is used to test for economies of scale. Cumulative industry output is used to test for firms' efficiency and learning effects. We regress the average prices on a constant, cumulative industry output, current industry output, and a set of dummy variables for different generations. Table 4 shows the results when we formulate learning effects to be identical across generations. We are able to use more than 500 observations and get R squares higher than 80%. We perform OLS and 2 SLS regressions, in which we instrument for the current industry output by using the price for material, which is the world market price of silicon compiled by Metal Bulletin. We also use summary statistics from the supply side such as the number of firms in the market. A negative sign for the cumulative industry output is consistent with learning by doing. The negative sign on current industry output relates to increasing economies of scale in the industry. Table 5 shows the results when we allow for different learning effects across different generations. As shown in the table the learning effects are around 0.4. They are very comparable across different generations until the 256MB generation. The learning effects are lower for the 512MB and the 1 GB generation. The lower learning effects are probably due to the very few observations for the last two generations. To summarize, it is interesting to note at this point that learning effects do not differ much across different generations.

We also use patent data in order to approximate the innovative pace over time and use the NBER patent database established by Hall, Jafee, and Trajtenberg (2001). The patent database includes patents that were applied for and subsequently granted in the U.S between 1963 and 2002. We use U.S. patents because the U.S. is the world's largest technology marketplace and it has become routine for non-U.S.-based firms to patent in the U.S., see also Albert (1991). The database holds detailed information on approximately 3 million U. S. utility patents. The patent data themselves were procured from the Patent Office. We identified the patents that each DRAM producer holds in the DRAM market. Table 1a shows the number of patents for every time period between 1989 and 1997. The number of patent applications is around 143,109 patents per year (as of 1997). The number of patents in the semiconductor industry represents a share of about 5% to 10% of the patents overall. The number of DRAM patents represents a share of about 5% of the semiconductor patents. Our data confirm that the DRAM industry is characterized by a high degree of innovation.

3 Dynamic Oligopoly Model

This section outlines a model of dynamic competition between oligopolistic firms in the DRAM industry. The model is formulated as a state game model. A firm's action in a given period influences not only its own and rival firms' current profits, but also its own and rival firms' future states. Besides market demand and market structure, an important state that affects current and future profits is a firm's cost structure.

The cost structure depends on produced output, input prices, a firm's experience in the production process, and on its productivity. Experience is determined by learning-by-doing and spillovers. The first component is usually modelled as own cumulative past output, and the second component is usually modelled as other firms' cumulative past output or cumulative past industry output. A firm's output decision is therefore an investment into experience and influences its own and rival firms' cost structure.

We use a discrete-time infinite horizon model with time indexed by $t = 0, 1, \ldots, \infty$. There are I firms denoted by $i = 1, \ldots, I$. The set of firms includes potential entrants and incumbent firms. In each period, each firm i earns profits equal to $\pi_{it} = \pi(q_{it}, q_{-it}, s_t, v_{it})$, which are a function of own actions q_{it}, other firms' actions q_{-it}, a vector of state variables s_t describing the market conditions and a private shock v_{it} describing a firm's productivity which shifts marginal costs.

Relevant state variables are market demand d_t, input prices m_t, the set of producing firms n_t and a firm i's experience ex_{it}, i.e. $s_t = (d_t, m_t, n_t, ex_{it})$. Market demand d_t and input prices m_t are determined by a common shock. The number of firms in the market n_t is determined by the exit decision of incumbents and the entry decision of potential entrants. Incumbent firms decide whether to stay in the market and produce q_{it} or to exit and receive a fixed scrap value κ. Potential entrants decide whether to enter the market and to produce output q_{it} or to stay out of the market and produce no output. A firm i's experience ex_{it} has two components. The first component is a firm's proprietary experience x_{it} and the second component is spillovers x_{-it} that firm i receives from other firms. A firm i's proprietary experience x_{it} is its own cumulative past output, such that $x_{it} = \sum_{\tau=1}^{t-1} q_{i\tau}$. Or expressed differently, $x_{it} = x_{it-1} + q_{it-1}$ with $x_{i0} = 0$, where we assume there is no proprietary experience in the beginning of the product cycle. A firm i's spillovers x_{-it} are other firms' cumulative past output, such that $x_{-it} = \sum_{\tau=1}^{t-2} \sum_{j \neq i} q_{j\tau}$. Or again expressed differently, $x_{-it} = x_{-it-1} + \sum_{j \neq i} q_{-jt-2}$ with $x_{-i0} = 0$, where we assume there are no spillovers in the beginning of the product cycle. Potential entrants have no experience and receive no spillovers.

Before firms simultaneously set their action by choosing their output q_{it}, each firm i observes a private shock v_{it}, independently drawn from a distribution $G_i(.|s_t)$. The private shock may derive from variability in production costs, c_{it}. Firms productivity is modeled as a first order autoregressive process $\omega_{it} = \rho\omega_{it-1} + v_{it}$, where ρ is the persistence or autocorrelation parameter. The autocorrelation reflects the fact that firms that are more productive today are more likely to be more productive tomorrow. Since the firm's productivity is correlated over time, it represents our serially correlated unobserved state variable.

Each potential entrant additionally observes a shock $u_{i\tau}$, independently drawn from a distribution $H_i(.|s_{\tau_i})$, where τ_i is the period firm i enters the market. As entering firms immediately start to produce, this means that a firm that enters the market observes two private shocks. As the shocks are private information firms solve for Bayesian Nash equilibria.

Each firm i maximizes its future discounted payoffs conditional on the

initial state s_0, the initial value of private shock v_{i0} and the initial value of sunk cost u_{i0}:

$$\mathrm{E}_{v,u} \sum_{t=0}^{\infty} \beta^t [\pi_i(q_{it}, q_{-it}, s_t, v_{it}, u_{it}) | s_0, v_{i0}, u_{i0}] \tag{1}$$

where $\beta \in (0, 1)$ is the discount factor, which is set equal to 0.9.

3.1 Profits in the Product Market

A firm i's per period profits in the product market are sales minus cost

$$\pi_{it}(q_{it}, q_{-it}, s_t, v_{it}) = p(q_t, z_t, d_t)q_{it} - c(q_{it}, m_t, x_{it}, x_{-it}, \omega_{it-1}, v_{it})q_{it} - f_i \tag{2}$$

where $p(q_t, z_t, d_t)$ is the industry price as a function of the industry output $q_t = \sum_{i=1}^{n_t} q_{it}$, observable demand shifters z_t and a random shock d_t. $c(q_{it}, m_t, x_{it}, x_{-it}, \omega_{it-1}, v_{it})$ is firm i's marginal cost as a function of its output q_{it}, input prices m_{it}, proprietary experience x_{it}, spillovers x_{-it}, unobserved state ω_{it-1} and firm i's private shock v_{it}. f_i is firm i's fixed cost. We specify the inverse demand function p_t as follows:

$$p_t(q_t, z_t, d_t) = d_t q_t^{\delta_1} z_t^{\delta_z}, \tag{3}$$

where δ_1, the elasticity of the inverse demand, and δ_z are coefficients to be estimated. We assume there is no firm specific uncertainty about demand as this would not be identified from a private shock in marginal cost. We specify a firm i's marginal costs as a linear function of its arguments:

$$c_i(q_{it}, m_t, x_{it}, x_{-it}, \omega_{it-1}, v_{it}) = \theta_0 + \theta_1 q_{it} + \theta_2 m_t + \theta_3 x_{it} + \theta_4 x_{-it} + \rho\omega_{it-1} + v_{it}, \tag{4}$$

where we denote the vectors of coefficients with θ and ρ and v_{it} is drawn from a standard normal distribution. The initial condition for ω_i is derived from the fact that firms do not produce output q_i before the product cycle starts.

3.2 Entry and Exit Cost

A potential entrant incurs entry cost when it enters the product market and its profits in the first period of market appearance are

$$\pi_i(q_{i\tau_i}, q_{-i\tau_i}, s_{\tau_i}, v_{i\tau_i}, u_i) = p(q_{\tau_i}, z_{\tau_i}, d_{\tau_i})q_{i\tau_i} - c(q_{i\tau_i}, m_{\tau_i}, v_{i\tau_i})q_{i\tau_i} - f_i - u_{i\tau_i}, \tag{5}$$

where τ_i is the period firm i enters the market and u_i is the privately observed random shock before entering the market. Learning-by-doing x_i, spillovers x_{-i}, the unobserved state ω_i are equal to zero at the time of entering the market. The profits of an incumbent firm that leaves the market are

$$\begin{aligned}\pi_i(q_{iT_i}, q_{-iT_i}, s_{T_i}, v_{iT_i}) = {} & p(q_{T_i}, z_{T_i}, d_{T_i})q_{iT_i} - c(q_{iT_i}, m_{T_i}, x_{iT_i}, x_{-iT_i}, \omega_{iT_i-1}, v_{iT_i})q_{iT_i} \\ & - f_i + k\end{aligned} \tag{6}$$

where T_i is the period firm i leaves the market and k is the scrap value.

3.3 Transition of States

For a complete description of the state game, the transition between states has to be defined. Our state variable market demand d_t is determined by a common period-specific shock and therefore does not require any further assumptions on state transitions over time. However, our state variables experience x_{it} and spillovers x_{-it} are influenced by past actions. The laws of motion for those state variables is deterministic and described by its cumulative past own output

$$x_{it+1} = x_{it} + q_{it} \tag{7}$$

and the cumulative past output of other firms

$$x_{-it+1} = x_{-it} + \sum_{j \neq i} q_{jt-1}. \tag{8}$$

For (7) and (8), the initial condition is that the respective state is equal to zero. There is no output production before the product cycle starts and no experience and no spillovers at the begin of the product cycle.

This leaves us to define the transition of the number of firms in the market n_t from time t to time $t+1$. The number of firms in the market n_{t+1} is

$$n_{t+1} = n_t + ne_t - nx_t, \tag{9}$$

where ne_t is the number of entering firms and nx_t the number of exiting firms. The number of entering firms ne_t depends on the distribution of u_i. A firm i enters, when future expected profits are positive. The number of entering firms nx_t depends on the scrap value k. A firm i exits, when future expected profits are lower than the scrap value which is fixed but could be estimated in the second stage.

3.4 Firms' Strategies

Firms use Markov strategies $q_{it} = \sigma_i(s_t, v_{it})$, i.e. a firm's output q_{it} is a function of the state variables s_t and the private shock v_{it}, generating Markov-perfect Nash equilibrium. Rival firms' strategies are denoted by $q_{-it} = \sigma_{-i}(s_t, v_{-it})$. If behavior is given by a Markov strategy profile $\sigma = (\sigma_i(s_t, v_{it}), \sigma_{-i}(s_t, v_{-it}))$, firm i's expected profits given the state variables s_t can be written recursively:

$$\begin{aligned} V_i(s_t;\sigma) = \mathrm{E}_{v,u}[&\pi_i(\sigma_i(s_t, v_{it}, u_i), \sigma_{-i}(s_t, v_{-it}, u_{-i}), s_t, v_{it}, u_i) \\ &+\beta \int V_i(s_{t+1};\sigma) dP(s_{t+1}|\sigma_i(s_t, v_{it}, u_i), \sigma_{-i}(s_t, v_{-it}, u_{-i}), s_t, v_{it}, u_i)|s_t], \end{aligned} \tag{10}$$

where $V_i(s_t;\sigma)$ is firm i's ex-ante value function. A strategy profile σ is a Markov perfect equilibria if, given the strategy profile of rival firms $\sigma_{-i}(s_t, v_{-it}, u_{-i})$, firm i does not want to deviate from its strategy profile $\sigma_i(s_t, v_{it}, u_i)$, i.e.

$$V_i(s_t;\sigma) \geq V_i(s_t;\sigma_i\prime, \sigma_{-i}), \tag{11}$$

where $\sigma_i\prime$ is an alternative strategy for firm i.

The structural parameters of our model are the discount parameter β, the profit functions $\pi_1, \ldots, \pi_I$, the distribution of private shocks G and H following a standard normal distribution. To obtain estimates of these parameters, we build on the estimation method developed by Bajari, Benkard and Levin (2007). This is a two-stage procedure. The first stage includes the estimation of the policy function σ_i, the transition probabilities P and the value functions V_i. The second stage estimates the profit function π_i and the distribution G_i. We assume that a firm's productivity is unobserved. We therefore extend their estimation method to allow for unobserved state variables and recover the unobserved state by assuming that productivity is a monotonic function of learning-by-doing and past cumulative output of other firms.

4 Econometric Model

In the following we will present the econometric model. As mentioned above we follow the 2-stage estimation routine by Bajari, Benkard and Levin (2007). In the first stage, we estimate the policy functions and the value function. The second stage assumes that the policy function and the transition probabilities are parameterized by a finite vector, and that this vector can be consistently estimated at the first stage. This assumption permits a non-parametric first stage with discrete action and state variables or a parametric first stage with continuous action and state variables. As described above, our model allows for continuous action and state variables. To parameterize the first stage, we thus have to assume that the functional form of the policy function and the transition probabilities is known. For the exposition of the estimation algorithm, we assume it is a linear function. The estimation algorithm is however equally applicable to more complicated functions of however known form. For the estimations, we try various higher order polynomials to approximate an arbitrary non-linear policy function and finally use the specification with the highest fit. Since some of the generations are not long enough in the market to generate a sufficiently large time series, we will not estimate the dynamic model for each generation separately, but will rather pool the data and use dummy variables to accounting for generation-specific effects. We also would like to refer to our estimation results displayed in Table 4 and 5, which provide support that learning effects are very comparable throughout different generations.

4.1 Estimation of the First Stage

The estimation of the first stage includes the estimation of the policy functions. There are various policy decisions. There is the entry decision of potential entrants and incumbents' decision whether to exit the market. Incumbents also decide on their output. For the incumbents' output function, it is necessary to obtain estimates for the demand (3) and entry profits (5).

Demand We specify the demand function log-linearly as

$$\ln(q_t) = \delta_0 + \delta_1 \ln(p_t) + \delta_2 \ln(p_t^S) + \delta_3 \ln(GGDP_t) + \delta_4 time_t + \sum_{l=5}^{15} \delta_l D_l + d_t \tag{12}$$

where we denote the vector of coefficients with δ. q_t is the market output of the chip at time t. p_t is the average selling price of a chip at time t, and p_t^S is the average selling price of the closest substitute. For the price of the closest substitute we construct a price index. For each DRAM generation, we identify a corresponding SRAM generation and use the price of this generation as the price of the closest substitute. $GGDP_t$ represents the growth rate of the GDP, which we use as an exogenous demand shifter. $Time$ is a time trend, D_l represents a dummy variables for every generation, where the 4K generation is used as the reference. d is a sequence of independently distributed normal variables with a mean of zero and a constant variance σ_d. We predict a negative sign for the own price elasticity of demand δ_1. The cross-price elasticity δ_2 is supposed to be positive (negative) if the respective products are substitutes (complements). We further await a positive sign for the demand shifter δ_3. The expected sign of the time trend coefficient δ_7 is supposed to be negative. It captures the effect of the time length that a particular generation has been in the market.

Incumbents' Output Policy Function Firm i's policy function σ_i is a function of the state variables s_t and the private shock v_{it} in marginal cost, i.e. $q_{it} = \sigma_i(s_t, v_{it}, u_i)$. If we assume that the policy function is log-linear in the state variables and in the private shock and if we implement the first order autoregressive process of the firm-level productivity, the policy function of incumbent firms is equal to

$$\begin{aligned} \ln(q_{it}) = {} & \gamma_0 + \gamma_1 \hat{d}_t + \gamma_2 \ln(m_t) + \gamma_3 \ln(n_t) + \gamma_4 \ln(x_{it}) + \gamma_5 \ln(x_{-it}) + \gamma_6 time_t \\ & + \gamma_7 w_{it-1} + \sum_{l=8}^{18} \gamma_l D_l + v_{it}, \end{aligned} \tag{13}$$

where we denote the vector of coefficients with γ, q_{it} represents firm i's output at time t and $\hat{d}_t$ is the contemporary demand shock obtained as the residual of (12). The variable m_t represents the price of silicon in period t, n_t stands for the lagged number of firms, x_{it} and x_{-it} represents the cumulative past output of firm i and all other firms, respectively. The *time* variable and the dummy variables are defined as in the demand equation. Note that we estimate a pooled regression in order to be able to use more observations for our variables of interest. Therefore, we assume that our right hand side variables have an equal impact on different generations. The dummy variables, however, will absorb any time invariant differences between the generations. Finally, w_{it-1} account for first order autocorrelation effects in firm i's production. We assume that a firm i's private shock v_{it} in marginal cost is uncorrelated with the state variables $s_t, s_{t-1}, \ldots, s_0$ such that

$$E[v_{it}|s_t, s_{t-1}, \ldots, s_0] = 0.$$

We would expect positive signs for the coefficients γ_2, γ_4, and γ_5 and a negative sign for γ_7.

We alsodirectly control for the serially correlated unobserved productivity by applying a lagged dependent variable model, or an AR(1) model. In this case the policy function looks as follows:

$$\begin{aligned}\ln(q_{it}) = {} & \widetilde{\gamma}_0 + \widetilde{\gamma}_1 \hat{d}_t + \widetilde{\gamma}_2 \ln(m_t) + \widetilde{\gamma}_3 \ln(n_t) + \widetilde{\gamma}_4 \ln(x_{it}) + \widetilde{\gamma}_5 \ln(x_{-it}) + \widetilde{\gamma}_6 time_t \\ & + \widetilde{\gamma}_7 \ln(q_{it-1}) + \sum_{l=8}^{18} \widetilde{\gamma}_l D_l + c_i + v_{it},\end{aligned} \tag{14}$$

where the vector of coefficients is denoted by $\widetilde{\gamma}$, and c_i denotes firm invariant unobserved heterogeneity. Finally, we can write the policy function in first differences in order to eliminate the unobserved heterogeneity.

Entry and Exit To obtain estimates for the distribution of u_i and κ, we estimate probit models. Potential entrants make their decision to enter dependent on the state variables d_t and n_t, but not on x_{it} and x_{-it} as they have not gained either propriety experience or gained experience through spillovers. We, however, augment the empirical model with cumulative industry output to account for cost efficiencies in the industry

$$\begin{aligned}P(\text{entry}_{\tau_i}) = {} & \alpha_0 + \alpha_1 \hat{d}_{\tau_i} + \alpha_2 \ln(m_{\tau_i}) + \alpha_3 \ln(n_{\tau_i}) + \alpha_4 \ln(x_{i\tau_i}) + \alpha_5 \ln(x_{-i\tau_i}) \\ & + \alpha_6 time_{\tau_i} + \sum_{l=7}^{17} \alpha_l D_l + u_{i\tau_i},\end{aligned} \tag{15}$$

where we denote the vector of coefficients with α, and $\hat{d}$ is the demand shock obtained as the residual of (12). All other variables are redundant to the previous specifications.

Incumbent firms face the decision, whether to stay in the market or to exit. Their decision to exit the market depends on all state variables

$$P(\text{exit}_{T_i}) = \lambda_0 + \lambda_1 \hat{d}_{T_i} + \lambda_2 \widehat{v}_{i\tau_i} + \lambda_3 \ln(m_{T_i}) + \lambda_4 \ln(n_{T_i}) + \lambda_5 \ln(x_{iT_i}) \qquad (16)$$
$$+\lambda_6 \ln(x_{-iT_i}) + \lambda_7 time_{T_i} + \sum_{l=8}^{18} \lambda_l D_l + \kappa_{T_i},$$

where we denote the vector of coefficients with λ and $\widehat{v}$ is the productivity shock obtained as the residual of the output policy function. Given (15) and (16), we calculate the number of firms in the market.

Value Functions Estimation of the value functions is based on the estimated policy functions and the transition between states. From estimating (13), we obviously get $q_{it} = \hat{q}_{it} + v_{it}$, which we use to simulate a sample of optimal policies

$$q_{itl} = \hat{q}_{it} + v_{itl}, \qquad (17)$$

where at each point in time $t = 0, 1, \ldots$, we draw a random sample of v_{itl} with $l = 1, \ldots, L$ from the distribution $G_i(.|s_t)$ and calculate simulated profits $\pi_{ilt}(q_{itl}, q_{-itl}, s_{tl}, v_{itl})$. We use (9) to move from one state to the other w.r.t. the number of firms and obtain for each simulation l

$$n_{t+1l} = n_{tl} + ne_{tl} - nx_{tl}, \qquad (18)$$

where ne_{tl} and nx_{tl} are determined by (15) and (16) and a random draw u_{itl} from $H_i(.|s_t)$ with $l = 1, \ldots, L$. We then use (7) and (8) to move from one state to the other w.r.t. proprietary experience and spillovers and obtain for each simulation l, $x_{it+1l} = x_{itl} + q_{itl}$ and $x_{-it+1l} = x_{-itl} + \sum_{j \neq i} q_{jt-1l}$. Finally, we use the specifications for demand (3) and the marginal cost function (4) and calculate simulated profits as

$$\pi_{itl} = \hat{p}_{tl} q_{itl} - (\theta_0 + \theta_1 q_{itl} + \theta_2 m_{tl} + \theta_3 x_{itl} + \theta_4 x_{-itl} + \rho\omega_{it-1l} + v_{itl})\, q_{itl} - f_i, \qquad (19)$$

where $\hat{\delta}_1$ is an estimate for the elasticity of demand obtained from (12) and $\hat{\omega}(x_{itl}, x_{-itl})$ is an estimate for $\omega(x_{itl}, x_{-itl})$ obtained from (13). To obtain an estimate for the value function, we add up profits π_{itl} over t and take the mean of the simulated profits π_{il} over l

$$\tilde{V}_i(s_t; \sigma_i, \sigma_{-i}, \delta, \gamma, \alpha, \lambda, \theta) = \qquad (20)$$
$$\frac{1}{L} \sum_{l=1}^{L} \sum_{t=0}^{\infty} \beta^t \{\hat{p}_{tl}\, q_{itl} - (\theta_0 + \theta_1 q_{itl} + \theta_2 m_{tl} + \theta_3 x_{itl} + \theta_4 x_{-itl} + \rho\omega_{it-1l} + v_{itl})\, q_{itl} - f_i\},$$

where we assume that for large enough t firms do not produce anymore.

4.2 Estimation of the Second Stage

To recover the structural parameters θ of the marginal cost function, we exploit the equilibrium condition (11) and construct for each simulated policy (17) an alternative policy and compare simulated and alternative strategies. An alternative strategy is equal to

$$q_{itl\prime} = q_{itl} + \epsilon, \tag{21}$$

where ϵ is a random draw from the standard normal distribution function. We now calculate alternative profits given alternative strategy $q_{itl\prime}$

$$\pi_{itl\prime} = \hat{p}_{tl\prime}\, q_{itl\prime} - (\theta_0 + \theta_1 q_{itl\prime} + \theta_2 m_{tl} + \theta_3 x_{itl\prime} + \theta_4 x_{-itl\prime} + \rho\omega_{it-1l\prime} + v_{itl}) q_{itl\prime} - f_i. \tag{22}$$

An estimate for the value function is again the mean of the simulated profits,

$$\tilde{V}_i(s_t; \sigma_{i\prime}, \sigma_{-i}, \delta, \gamma, \alpha, \lambda, \theta) =$$
$$\frac{1}{L}\sum_{l=1}^{L}\sum_{t=0}^{\infty}\beta^t\{\hat{p}_{tl\prime}\, q_{itl\prime} - (\theta_0 + \theta_1 q_{itl\prime} + \theta_2 m_{tl} + \theta_3 x_{itl\prime} + \theta_4 x_{-itl\prime} + \rho\omega_{it-1l\prime} + v_{itl}) q_{itl\prime} - f_i\}.$$

For ϵ we use $m = 1, \ldots, M$ random numbers, ϵ_m. This gives us $M \times \tilde{V}_i(s; \sigma_{i\prime}, \sigma_{-i}, \delta, \gamma, \alpha, \lambda, \theta)$'s. We can rewrite the equilibrium condition (11) as

$$\tilde{V}_i(s_t; \sigma_i, \sigma_{-i}, \delta, \gamma, \alpha, \lambda, \theta) \geq \tilde{V}_i(s_t; \sigma_{i\prime}, \sigma_{-i}, \delta, \gamma, \alpha, \lambda, \theta). \tag{23}$$

Exploiting the linearity of θ in firm i's profit, provides us with

$$[\tilde{W}_i(s_t; \sigma_i, \sigma_{-i}, \delta, \gamma, \alpha, \lambda) - \tilde{W}_i(s_t; \sigma_{i\prime}, \sigma_{-i}, \delta, \gamma, \alpha, \lambda)]\ \theta \geq 0. \tag{24}$$

Let's define a function g as follows

$$g(y; \delta, \gamma, \alpha, \lambda, \theta) := [\tilde{W}_i(s_t; \sigma_i, \sigma_{-i}, \delta, \gamma, \alpha, \lambda) - \tilde{W}_i(s_t; \sigma_{i\prime}, \sigma_{-i}, \delta, \gamma, \alpha, \lambda)]\ \theta \geq 0. \tag{25}$$

The inequality defined by y is satisfied at $(\delta, \gamma, \alpha, \lambda, \theta)$, if $g(y; \delta, \gamma, \alpha, \lambda, \theta) \geq 0$. We can express (25) also in differences,

$$g_{itl}(y_{itl}; \delta, \gamma, \alpha, \lambda, \theta) := \tag{26}$$
$$(\hat{p}_{tl} - \hat{p}_{tl\prime})\Delta q_{itl} - (\theta_0 + \theta_1 \Delta q_{itl} + \theta_2 m_{tl} + \theta_3 \Delta x_{itl} + \theta_4 \Delta x_{-itl} + \rho\Delta\omega_{it-1l} + \Delta v_{itl})\Delta q_{itl}$$

where $\Delta q_{itl} = q_{itl} - q_{itl\prime}$. When we define the function

$$Q(\delta, \gamma, \alpha, \lambda, \theta) := \int (\min\{g(y; \delta, \gamma, \alpha, \lambda, \theta), 0\})^2 dH(x). \tag{27}$$

When we define the function $\tilde{g}(y;\hat{\delta},\hat{\gamma},\hat{\alpha},\hat{\lambda},\theta)$ as the empirical counterpart of $g(x;\delta,\gamma,\alpha,\lambda,\theta)$ computed by replacing the V_i terms with simulated estimates $\tilde{V}_i$, we can define

$$Q_n(\delta,\gamma,\alpha,\lambda,\theta) := \sum_{l=1}^{L}\{\min\,[\tilde{g}(y;\hat{\delta},\hat{\gamma},\hat{\alpha},\hat{\lambda},\theta),0]\}^2. \tag{28}$$

Finally, in order to estimate the sunk cost we rely on the fact that there firms are forward-looking and rational and are able to calculate their discounted profit stream given the evolution of the state vector and their policy functions. If a firm does not enter, even though the expected profits are positive, it implies that the draw on sunk cost exceeded the value generated in the market. Hence, the discounted expected profits of entering at different states is simulated for many different paths. Averaging those gives the theoretically expected profits of entering at different states. The distance between those caclulated profit streams and the observed observed entry rates at those states is minimized, which gives allows us to recover the sunk cost distribution.

Finally, we calculate the sunk costs by calculating the expect discounted values at different states and comparing them to entry observations at those states. If entry occured at those states it indicates that sunk costs are lower than the generated discounted profits at this stage and vice versa.

5 The Estimation Results

This section discusses the estimation results. We start with the estimation results of the demand function. We then proceed with the entry and exit distribution, and the incumbents' output policy function. Finally, we describe the structural parameters.

Demand Equation In order to obtain estimates for the coefficient vector δ, we estimate industry demand (12) using ordinary least squares as well as two stage least squares. In the latter case, we instrument the average selling price in the demand equation using the GDP of the OECD at constant prices as a nonprice demand shifter. We also use summary measures from the supply side, like the number of firms in the industry, cumulative industry output, and the price of silicon – all variables in logarithm.

The estimation results of the demand equation are shown in Table 6. The results using the ordinary least squares estimator as shown in column 1, whereas the results for the 2 stage least squares estimator as depicted in columns 2 and 3. Since the results of the two estimators are very similar, we will exclusively report on the results using the instrumental variable estimator.

The first stage equation (column 2) represents a good fit with an adjusted R-square of about 0.94. A test for the joint significance of the instruments indicates that the number of firms in the industry, cumulative world output and the price of silicon are highly correlated with the average selling price. With a value of 73.41 for the F-statistics, we reject the null hypothesis that the estimated coefficients of these variables are equal to zero. A Hausman test indicates the necessity to instrument the average selling price in the demand equation. The value of the χ^2 distributed test statistic is equal to 60.55, which is larger than 18.31 – the 5% critical value with 11 degrees of freedom. Two of our three instruments are significantly different from zero. The negative sign on the cumulative industry output is meaningful as higher cumulative industry output lowers marginal costs in the presence of learning by doing, which shifts the supply curve downwards resulting in lower equilibrium prices. The positive sign on the price of silicon indicates that higher factor prices shifts the marginal cost curve upwards which results in higher equilibrium prices.

Turning to the second stage of the instrumental variable estimator (column 3), the R-square of about 0.71 confirms a high explanatory power of the estimation. All variables are significantly different from zero at least at the 5% level. The estimate of the average selling price of a chip is a negatively significant indicating a negative own price elasticity of demand. The magnitude of -3.03 represents the fact the DRAM market is characterized by a highly elastic demand curve. The estimate of SRAM chips, is significant and positive, indicating a positive cross-price elasticity and indicating that SRAM chips represent substitutes. Moreover, the estimate of 0.458 also confirms that the price of SRAMs has a lower impact on the DRAM demand than the price of DRAMs themselves. The demand shifter $GGDP_t$ is positively significant, providing evidence that a higher growth in GDP shifts the demand outwards. The negative time trend is consistent with previous findings that buyers substitute away from one generation to the next as time elapses. The dummy variables for the different generations are all highly significant and positive. The magnitude of the estimates are very comparable over different generations, which reinforces the reliability of the estimation results. Interestingly, the magnitude of the dummy variables is increasing throughout all the different generations, which underlines the increasing importance of using DRAM chips in application specific electronic products. Moreover, it is interesting to note that the increase in dummy variables increasees by a magnitude of 3 to 4 up until 16 MB generation. Thereafter, however, the increase in the dummy variables diminishes to 1. This results emphasizes that the growth in market demand increased over different generations, but the growth in demand slowed down towards the more recent generations.

Policy Function We estimate incumbents' output policy (13) with general least squares to obtain estimates for the coefficient vector γ. The results are shown in Table 7. We estimate equation (13) in by accounting for a first-order autocorrelation process, see column 1. We also estimate firm's output policy function by applying a lagged dependent variable model, or an AR(1) model, in order to control for the serially correlated unobserved productivity as shown in (14). Finally, we apply a first difference estimator accounting for a first-order autocorrelation process in the unobserved state variable ω_i (columns 3 and 4, respectively). Column 3 display the estimation results for the Arellano-Bond (1991) estimator which uses lagged dependent variables in levels. Column 4 displays the results for the Blundell-Bond (1998) estimator, which uses the levels and differences of the lagged dependent variable in the set of instruments. Note that the last two estimators eliminate the unobserved heterogeneities by applying first differences.

Our pooled regression allows us to use approximately 3,500 observations. The regression estimations for the instrumental variable estimations illustrates a remarkably good fit, it has R-square of 0.89. The estimates for the instrumental variable regressions in differences performs quite poorly and also carries parameter estimates that are sometimes couterintuitive. The problem with the first difference estimators is that the instruments are not strongy correlated as the series on production is highly persistent, so that lagged levels are only weakly correlated with first differences. The instrumental variable estimator in levels (column 2) fits the expectations of our model the best. The observed serially correlated variable, cumulative past output, which captures the learning by doing effects is positive and significant. This result emphasizes the importance of learning by doing in this industry. More experience in production increases efficiency and increases output. The lagged output carries a positive significant sign which shows that a first-order autocorrelation process is present in the data. We can confirm that correcting for serially correlated unobserved state variable, e.g. firm-specific productivity is important to control for. It confirms our notion that firms are able to react according to the private shocks they receive in the short run. The positive demand shock indicates that firms are able to increase their production. The negative sign on the price of material confirms that higher factor prices increase marginal cost and lower firm level output. The positive sign of the number of firms in the market illustrates that more firms in the market increase the competitive pressure in the market. The dummy variables for the different generations as well as the time trend turn out to be highly significant.

Entry and Exit Distribution We estimate the entry distribution (15) and exit distribution (16) with probit models to obtain estimates for the coefficient vectors α and λ. The results are shown in Table 8. The results for the entry regression are shown in the first column.

The positive coefficient on number of firms is insignificant. However, the sign indicates the fact that few early movers enter at the initial time periods, whereas the majority of firms enters when the life cylce approaches the matured phases. This entry pattern emphasizes the fact that firms need to come up with a new technology to enter a new technology, and only few firms are clearly ahead of others. This results reinforces the existence of spillovers in the market, which make it difficult to protect flows from research and development. The time trend shows that the number of entering firms increases over the life cycle of a generation, which is intiutive as we include the whole time span over the life cycle for most generations. An interesting results is that the dummy variables are negative and become even more negative throughout different generations. This result shows that entry became less likely over different generations given generation-specific fixed effects, which is an indication that entry costs increased throughout different generations. An increase in sunk costs would be supported by the engineering literature. Slightly surprising is the negative sign of the demand shock.and the positive coefficient on the price of silicon.

Turning to the results for the exit equation (columns 2-4), the demand and productivity shock carry negative signs. The results confirm that negative productivity shocks foster firm's exit.

Structural Parameters Finally, we are interested in recovering the structural parameters θ and ρ from the marginal cost function (4) as well as estimating the sunk cost in the different generations. As described above, we exploit the equilibrium condition (11) and construct for each simulated policy (17) an alternative policy. We compared the simulated value functions based on optimal strategies with the simulated values based on alternative non-optimal strategies and minimize the deviations of those to recover the structural parameters. We use 10,000 simulations and without firm-specific and product-specific fixed effects in the marginal cost function.

As shown in Table 9, the estimation results are plausible. We are able to use 304 observations and our structural parameters are all highly significant at the 1% level. We find that the cost function is characterized by increasing economies of scale. Moreover, we can confirm significant learning by doing effects and spillovers being prevalent in the industry which lower the marginal costs. Our estimate for sunk costs over all generations are about 1.3 billion US-dollars and get close to what has been reported in business reports. The standard deviation is about 2.2 billion US-dollars, which indicates that sunk cost over different generations do fluctuate a lot. We can also confirm increasing sunk cost over the first part of the different generations. However, we are currently facing the problem that some estimated sunk costs for the latest generations are decreasing. We think that we may not have enough data to accurately estimate the sunk costs for the latest generations. The expected discounted profits are not accurate in this case as the life cycles did not even reach the peak yet. Since the discounted profits are compared to the entry probabilities we may get distorted sunk cost estimates. We contemplate to possibly correct for this truncation problem.

6 Conclusion

This study tried to investigate why the number of firms increased during early DRAM generations and the sharply declines throughout more recent DRAM generations. We estimated a dynamic model accounting for entry, exit, and intertemporal production decision made by firms in an oligopolistic market structure. We build on the estimator by Bajari, Benkard and Levin (2007) and accounted for additional serially correlated unobserved state variables. Those unobservables are supposed to capture unobserved productivity differences between firms that enter the marginal cost function.

We could also show that implementing an unobserved state variable is indeed important and turns out to return significant estimates. Our sunk costs estimates are getting close to the reported establishment plant costs.

We find that growth in market demand increased over different generations. However, we also find that the growth rates are higher for early generations. The growth rates slow down for the more recent generations. Entry costs increased throughout early generations. We can conclude that for early generations the growth in demand overcompensates the losses from increasing sunk costs. For more recent generations, however, the growth in demand diminishes and the sunk cost increases. Since the losses from increasing sunk costs dominates the benefits from diminishing growth rates, fewer firms enter the more recent generations. Hence, the inverse U-shape in the industry can be explained by the interdependence between the growth in market demand and the increase in sunk costs.

What is left to do:

We need to report several tests such as validity of instruments, and sequential exogenous regressors.

We also want to do more robustness checks with regard to the number of simulations.

We want to test how the number of firms over generations evolves once we hold the sunk costs constant over all generations. It would be interesting to see if the model generates an inverse U-shape of the number of firms even when sunk cost are set constant. If we don't get close to the data, we could confirm that the increase in sunk cost is important to generate the market structure that we observe in the data.

References

Ackerberg, D., L. Benkard, S. Berry, A. Pakes, 2006, "Econometric Outcomes for Analyzing Market Outcomes" in The Handbook of Econometrics, J. J. Heckman (ed.).

Ackerberg, D., K. Caves and G. Frazer, 2005, "Structural Identification of Production Functions", Mimeo, UCLA.

Aguirregabiria, V. and P. Mira, 2007, "Sequential Estimation of Dynamic Games", Econometrica, Vol. 75, No. 1, pp. 1-53.

Asplund, M. and V. Nocke, 2006, "Firm Turnover in Imperfectly Competitive Markets", Review of Economic Studies, Vol. 73, No. 2, pp. 295-327.

Audretsch, D. and S. Klepper, 2000, "Innovation, Evolution of Industry and Economic Growth", Mimeo, Indiana University.

Bajari, P., L. Benkard and J. Levin, 2007, "Estimating Dynamic Models of Imperfect Competition", Econometrica, Vol. 75, No. 5, pp. 1331–1370.

Baumol, W., J. Panzar, and R. Willig, 1982, "Contestable Markets and the Theory of Industry Structure", New York: Harcourt Brace Jovanovich.

Benkard, L., 2004, "A Dynamic Analysis of the Market for Wide-Bodied Commercial Aircraft", Review of Economic Studies, Vol. 71, No. 3, pp. 581-612.

Beresteanu, A. and B. Ellickson, 2006, "The Dynamics of Retail Oligopoly", Mimeo, Duke University.

Berndt, E. R., 1991,.*The Practice of Econometrics,* Addison Wesley, Reading Massachusetts.

Berry, S., J. Levinsohn and A. Pakes, 1995, "Automobile Prices in Market Equilibrium", Econometrica, Vol. 63, No. 4, pp. 841-890.

Bresnahan, T. F., and P.C. Reiss, 1989, "Do Entry Conditions Vary Across Markets", Brookings Papers on Economic Activity, Vol. 3, pp. 883-871.

Bresnahan, T. F., and P.C. Reiss, 1990a, "Entry in Monopoly Markets", Review of Economic Studies, Vol. 57, pp. 531-553.

Bresnahan T. and P.C. Reiss, 1991, "Empirical Models of Discrete Games", Journal of Econometrics, Vol. 48, pp. 57–81.

Bresnahan, T. and P.C. Reiss, 1991, "Entry and Competition in Concentrated Markets", Journal of Political Economy, Vol. 99, No. 5, pp. 977-1009.

Collard-Wexler, A., 2005, "Plant Turnover and Demand Fluctuations in the Ready-Mix Concrete Industry", Mimeo, Northwestern University.

Dataquest, 2008, *Semiconductor Database provided by Dataquest.*

Dick, A. R., 1991, "Learning by Doing and Dumping in the Semiconductor Industry", Journal of Law and Economics, Vol. 34, pp. 133-159.

Duflo, E., R. Hanna and S. P. Ryan, 2007, "Monitoring Works: Getting Teachers to Come to School", Mimeo, Massachusetts Institute of Technology.

Ericson, R. and A. Pakes, 1995, "Markov-Perfect Industry Dynamics: A Framework for Empirical Work", Review of Economic Studies, Vol. 62, No. 1, pp. 53-83.

Flamm, K., 1996, "Mismanaged Trade?", Brookings Institution Press, Washington, D.C.

Geroski, P. A., 1995, "What Do We Know About Entry?", International Journal of Industrial Organization, Vol. 13, pp. 421-440.

Gowrisankaran, G. and R. Town, 1997, "Dynamic Equilibrium in the Hospital Industry", Journal of Economics and Management Strategy, Vol. 6, No. 1, pp. 45-74.

Gruber, H., 1996, "Trade Policy and Learning by Doing: The Case of Semiconductors", Research Policy, Vol. 25, pp. 723-739.

Hotz, V., R. Miller, S. Sanders and J. Smith, 1993, "A Simulation Estimator for Dynamic Models of Discrete Choice", Review of Economic Studies, Vol. 60, pp. 397-429.

Irwin, D. and P. Klenow, 1994, "Learning by Doing Spillovers in the Semiconductor Industry", Journal of Political Economy, Vol. 102, pp. 1200-1227.

Jofre-Bonet, M. and M. Pesendorfer, 2003, "Estimation of a Dynamic Auction Game", Econometrica, Vol. 71, No. 5, pp. 1443-1489.

Klepper, S., and K. Simons, 1996, "Innovation and Industry Shakeouts", Mimeo, Carnegie Mellon University.

Levinsohn, J. and A. Petrin, 2003, "Estimating Production Functions Using Inputs to Control for Unobservables", The Review of Economic Studies, Vol. 70, No. 2, pp. 317-341.

Lieberman, M.B., 1984, "The Learning Curve and Pricing in the Chemical Processing Industries", Rand Journal of Economics, 15, pp. 213-228.

Macieira, J., 2006, "Measuring the Effects of Innovation Subsidies: The Case of Supercomputers", Mimeo, Northwestern University.

Mazzeo, M. J., 2002, "Product Choice and Oligopoly Market Structure", RAND Journal of Economics, Vol. 33, No. 2, pp. 221-242.

Mueller, D. C. and J. E. Tilton, 1969, "Research and Development Costs as a Barrier to Entry", Canadian Journal of Economics, Vol. 4, pp. 570-579.

OECD/IEA, 1998, "Energy Prices and Taxes, International Energy Agency, First Quarter to Fourth Quarter,1984-1998", Paris.

OECD, 1998, *STAN DATABASE*, Paris.

Olley, S. and A. Pakes, 1996, "The Dynamics of Productivity in the Telecommunications Equipment Industry", Econometrica, Vol. 64, No. 6, pp. 1263-1298.

Pakes, A. and P. McGuire, 1994, "Computing Markov-Perfect Nash Equilibria: Numerical Implications of a Dynamic Differentiated Product Model", RAND Journal of Economics, Vol. 25, No. 4, pp. 555-89.

Pakes, A. and P. McGuire, 2001, "Stochastic Approximation for Dynamic Models: Markov Perfect Equilibrium and the 'Curse' of Dimensionality", Econometrica, Vol. 69, No. 5, pp. 1261-81.

Pakes, A., M. Ostrovsky and S. Berry, 2006, "Simple Estimators for the Parameters of Discrete Dynamic Games (with Entry/Exit Examples)", RAND Journal of Economics, forthcoming.

Pesendorfer, M., 2003, "Horizontal Mergers in the Paper Industry", RAND Journal of Economics, Vol. 34, No. 3, pp. 495-515.

Pesendorfer, M. and P. Schmidt-Dengler, 2003, "Identification and Estimation of Dynamic Games", Working Paper #9726, NBER.

Pesendorfer, M. and P. Schmidt-Dengler, 2007, "Asymptotic Least Squares Estimators for Dynamic Games", Mimeo, London School of Economics.

Ryan, S., 2006, "The Costs of Environmental Regulation in a Concentrated Industry", Mimeo, Massachusetts Institute of Technology.

Seim, K., 2006, "An Empirical Model of Firm Entry with Endogenous Product-Type Choices", RAND Journal of Economics, Vol. 37, No. 3, pp. 619-640.

Siebert, R., 2007, "Learning by Doing and Multiproduction Effects over the Life Cycle: Evidence from the Semiconductor Industry", Mimeo, Purdue University.

Sweeting, A., 2006, "The Costs of Product Repositioning: The Case of Format Switching in the Commercial Radio Industry", Mimeo, Northwestern University.

Toivanen, O. and Waterson, M., 2005, "Market Structure and Entry: Where's the Beef?", RAND Journal of Economics, Vol. 36, pp. 680-699.

Wooldridge, J., 2005, "On Estimating Firm-Level Production Functions Using Proxy Variables to Control for Unobservables", Mimeo, Michigan State University.

Wright, T. P., 1936, "Factors Affecting the Cost of Airplanes", Journal of Aeronautical Sciences, Vol. 3, pp. 122-128.

Zulehner, C., 2003, "Testing Dynamic Oligopolistic Interaction: Evidence from the Semiconductor Industry", International Journal of Industrial Organization, Vol. 21, No. 10, pp. 1527-1556.

7 Appendix: Tables and Figures

Table 1a: Number of Patents

Time	DRAM Patents	Semiconductor Patents	Total Patents
1989	701	4,063	78,619
1990	765	4,521	81,302
1991	853	5,276	82,939
1992	908	5,313	86,548
1993	1,083	5,688	89,572
1994	1,361	7,554	102,553
1995	1,572	9,250	122,127
1996	1,951	10,390	122,552
1997	2,390	13,507	143,109

Table 1a presents number of patents over time. Source: NBER patent database.

Table 1b: Reported Establishment Costs

COMPANY	COUNTRY	PRODUCTS	YEAR	Wafer/Month	Cost (US$)
Fujitsu	England	64MB DRAM	1999	15,000	1.4b
IBM	France	16/64MB DRAM	1997	20,000	1.0b
Siemens	Germany	256MB DRAM	1999	25,000	1.9b
Siemens	England	Memory	1997	25,000	1.6b
Texas Instr	Italy	16MB DRAM	1995	15,000	1.0b
LG	Wales	256MB DRAM	1998	TBA	1.3b

Table 1b presents reported establishment costs. Prices are in current US Dollars. Souce: Gruber (199?).

Table 2a: Firms in the DRAM Industry

Firms	4K	16K	64K	256K	1Mb	4Mb	16Mb
Adv. Micro Dev.	x	x	x	.	.	.	.
Alliance	.	.	.	.	.	x	x
Am. Microsyst.	x	.	.	.	.	.	.
AT&T	.	.	.	x	x	.	.
Elpida	.	.	.	.	.	.	x
Eurotechnique	.	x	.	.	.	.	.
Fairchild	x	x	x	.	.	.	.
Fujitsu	x	x	x	x	x	x	x
G-Link	.	.	.	.	x	x	x
Hitachi	x	x	x	x	x	x	x
Hynix	.	.	.	.	.	x	x
Hyundai	.	.	x	x	x	x	x
IBM	.	.	.	.	x	x	x
Inmos	.	.	x	x	.	.	.
Intel	x	x	x	x	x	.	.
Intersil	x	x	.	.	.	.	.
LG	.	.	.	x	x	x	x
Matsushita	.	x	x	x	x	x	x
Micron	.	.	x	x	x	x	x
Mitsubishi	.	x	x	x	x	x	x
Mosel Vitelic	.	.	x	x	x	x	x
Mostek	x	x	x	x	.	.	.
Motorola	x	x	x	x	x	x	x
Nan Ya Techn.	.	.	.	.	.	.	x
Ntl. Semic.	x	x	x	x	.	.	.
NEC	x	x	x	x	x	x	x
Nippon	.	.	.	x	x	x	x
OKI	.	.	x	x	x	x	x
Ramtron Int.	.	.	.	.	.	x	.
Samsung	.	.	x	x	x	x	x
Sanyo	.	.	.	x	x	x	.
SGS-Thompson	x	x	.	.	.	.	.
Sharp	.	.	x	x	x	x	.
Siemens	.	x	x	x	x	x	x
Signetics	x	x	.	.	.	.	.
STC	x	x	x	.	.	.	.
Texas Instr.	x	x	x	x	x	x	x
Toshiba	.	x	x	x	x	x	x
Vanguard	.	.	.	.	.	x	x
Zilog	.	x	.	.	.	.	.
Number of Firms	**15**	**20**	**22**	**23**	**22**	**24**	**23**

Table 2a presents those firms which entered different DRAM generations. An x indicates a firm's presence in the DRAM market. Souce: Gartner.

Table 2b: Firms in the DRAM Industry

Firms	**64Mb**	**128MB**	**256MB**	**512MB**	**1GB**
Adv. Micro Dev.	.	.	.	.	.
Alliance	.	.	.	.	.
Am. Microsyst.	.	.	.	.	.
AT&T	.	.	.	.	.
Elpida	x	x	x	x	x
Eurotechnique	.	.	.	.	.
Fairchild	.	.	.	.	.
Fujitsu	x	x	.	.	.
G-Link	x	x	.	.	.
Hitachi	x	x	x	.	.
Hynix	x	x	x	x	x
Hyundai	x	x	.	.	.
IBM	x	.	.	.	.
Inmos	.	.	.	.	.
Intel	.	.	.	.	.
Intersil	.	.	.	.	.
LG	x	.	.	.	.
Matsushita	x	.	.	.	.
Micron	x	x	x	x	x
Mitsubishi	x	x	x	.	.
Mosel Vitelic	x	x	x	.	.
Mostek	.	.	.	.	.
Motorola	x	.	.	.	.
Nan Ya Techn.	x	x	x	x	.
Ntl. Semic.	.	.	.	.	.
NEC	x	x	.	.	.
Nippon	x	.	.	.	.
OKI	x	.	.	.	.
Ramtron Int.	.	.	.	.	.
Samsung	x	x	x	x	x
Sanyo	.	.	.	.	.
SGS-Thompson	.	.	.	.	.
Sharp	.	.	.	.	.
Siemens	x	x	x	x	x
Signetics	.	.	.	.	.
STC	.	.	.	.	.
Texas Instr.	x	.	.	.	.
Toshiba	x	x	x	x	.
Vanguard	x	x	x	.	.
Zilog	.	.	.	.	.
Number of Firms	**22**	**15**	**11**	**7**	**5**

Table 2b presents those firms which entered different DRAM generations. An x indicates a firm's presence in the DRAM market. Souce: Gartner.

Table 3: Summary Statistics

Variable	Number of observations	Mean	Standard Deviation	Minimum	Maximum
Raw variables					
Average selling price	311	26.878	59.739	1.081	542.038
Industry output $\times 10^3$	311	66451.13	89593.13	6	425000
Number of firms	311	12.724	5.518	1	23
Cumulative industry output	311	1346536	1564034	3	5234100
Price of silicon	311	1868.505	522.008	602.556	3533.678
GDP	311	1.42e+07	7087456	3593009	3.22e+07
Variables in logarithm					
Log(Price)	311	2.046	1.449	0.078	6.295
Log(Industry output $\times 10^3$)	311	9.476	2.469	1.792	12.959
Log(Number of firms)	311	2.399	0.610	0	3.136
Log(Cumulative industry output)	311	11.982	3.523	1.099	15.471
Log(Price of silicon)	311	7.496	0.274	6.401	8.17
Log(GDP)	311	16.443	0.166	16.045	16.673

Table 3 presents summary statistics. Prices are in constant US Dollars as of 1996. Several sources are mentioned in the text.

Table 4: Learning Effects in the DRAM Industry

Variable	OLS	OLS	IV
Constant	6.264	5.867	5.692
	(55.40)***	(52.93)***	(49.91)***
Log(Cumulative industry output)	-0.393	-0.467	-0.499
	(-53.10)***	(-46.43)***	(-46.83)***
Log(Output)		0.156	0.225
		(9.99)***	(13.14)***
Dummy variable for 16K	0.095	-0.175	-0.293
	(0.81)	(-1.58)	(-2.59)**
Dummy variable for 64K	0.150	0.147	0.146
	(1.48)	(1.58)	(1.54)
Dummy variable for 256K	0.526	0.346	0.267
	(4.98)***	(3.52)***	(2.66)**
Dummy variable for 1MB	0.902	0.629	0.509
	(8.48)***	(6.22)***	(4.91)***
Dummy variable for 2MB	-0.167	-0.380	-0.474
	(-1.28)	(-3.13)**	(-3.82)***
Dummy variable for 4MB	0.902	0.651	0.539
	(8.68)***	(6.61)***	(5.35)***
Dummy variable for 8MB	-0.289	-0.545	-0.658
	(-1.68)	(-3.41)***	(-4.03)***
Dummy variable for 16MB	1.127	0.783	0.631
	(10.47)***	(7.50)***	(5.88)***
Dummy variable for 64MB	0.848	0.451	0.275
	(7.05)***	(3.84)***	(2.28)*
Dummy variable for 128MB	0.718	0.269	0.072
	(5.47)***	(2.10)*	(0.54)
Dummy variable for 256MB	0.899	0.399	0.178
	(6.41)***	(2.89)**	(1.25)
Dummy variable for 256MB	0.899	0.399	0.178
	(6.41)***	(2.89)**	(1.25)
Dummy variable for 1GB	0.114	-0.098	-0.191
	(0.49)	(-0.45)	(-0.87)
Number of observations	526	526	526
R-squared adjusted	0.85	0.88	0.87

Table 4 presents learning effects (overall) for the DRAM industry. Absolute values of t-statistics are shown in parentheses below the parameter estimates, *** (**, *) denotes a 99% (95%, 90%) level of significance.

Table 5: Learning Effects in different DRAM Generations

Panel A.	16K	64K	256K	1MB	2MB
Constant	6.453	6.672	6.262	6.468	4.249
	(27.44)***	(51.01)***	(30.81)***	(21.02)***	(9.98)***
Log(Cumulative industry output)	-0.401	-0.413	-0.354	-0.341	-0.215
	(-20.09)***	(-41.36)***	(-24.16)***	(-15.15)***	(-5.40)***
Number of observations	37	68	60	57	26
R-squared adjusted	0.92	0.96	0.91	0.80	0.53

Panel B.	4MB	8MB	16MB	64MB	128MB
Constant	7.806	3.609	8.443	7.754	6.063
	(21.01)***	(17.51)***	(22.27)***	(18.74)***	(14.76)***
Log(Cumulative industry output)	-0.439	-0.174	-0.472	-0.441	-0.323
	(-16.66)***	(-9.13)***	(-17.09)***	(-14.67)***	(-10.53)***
Number of observations	65	12	54	35	26
R-squared adjusted	0.81	0.88	0.85	0.86	0.81

Panel C.	256MB	512MB	1GB
Constant	6.818	5.340	4.918
	(13.75)***	(32.04)***	(29.10)***
Log(Cumulative industry output)	-0.365	-0.250	-0.155
	(-9.40)***	(-13.38)***	(-5.96)**
Number of observations	21	12	6
R-squared adjusted	0.81	0.94	0.87

Table 5 presents different learning effects for different DRAM generations. Absolute values of t-statistics are shown in parentheses below the parameter estimates, *** (**, *) denotes a 99% (95%, 90%) level of significance.

Table 6: Results for the Demand Equation

	Ordinary least squares	Two-stage least squares	
		First stage	Second stage
Constant	14.057	-3.523	16.346
	(24.04)***	(-4.24)***	(25.84)***
Log(Average selling price)	-2.415		-3.031
	(-30.98)***		(-31.22)***
Log(Average SRAM selling price)	0.419	0.133	0.458
	(4.78)***	(4.81)***	(5.05)***
Time trend	-0.129	-0.007	-0.179
	(-14.48)***	(-1.79)	(-17.92)***
Log(GGDP)	36.749	10.351	45.589
	(2.56)*	(2.32)*	(3.06)**
Log(Number of firms)		-0.111	
		(-1.51)	
Log(Cumulative industry output)		-0.282	
		(-15.35)***	
Log(Price of silicon)		1.087	
		(11.05)***	
Dummy variable for 16K	2.564	0.174	2.675
	(8.45)***	(1.58)	(8.58)***
Dummy variable for 64K	3.943	0.528	5.126
	(12.49)***	(4.59)***	(15.21)***
Dummy variable for 256K	6.98	1.059	8.784
	(18.10)***	(6.45)***	(20.98)***
Dummy variable for 1MB	9.611	1.609	12.189
	(19.85)***	(7.75)***	(22.75)***
Dummy variable for 4MB	11.548	2.028	14.792
	(19.25)***	(8.04)***	(22.18)***
Dummy variable for 16MB	13.399	2.425	17.199
	(19.43)***	(8.32)***	(22.35)***
Dummy variable for 64MB	14.257	2.409	18.290
	(19.09)***	(7.65)***	(22.04)***
Dummy variable for 128MB	14.823	2.366	19.161
	(18.73)***	(7.37)***	(21.83)***
Dummy variable for 256MB	16.251	2.689	20.937
	(19.48)***	(8.06)***	(22.50)***
Dummy variable for 512MB	15.915	2.513	21.564
	(16.76)***	(6.49)***	(20.02)***
Dummy variable for 1GB	15.497	2.238	22.472
	(13.32)***	(4.47)***	(16.11)***
Number of observations	453	444	444
R-squared adjusted	0.76	0.93	0.71

Table 6 presents the ordinary least squares and two-stage least squares estimation results for the demand equation. In the demand equation (columns 1 and 3), the dependent variable is the industry output. Explanatory variables are the average selling price, a general demand shifter, and a time trend. In the reduced form supply equation (column 2), the dependent variable is the average selling price. Explanatory variables are the number of firms, cumulative industry output, and the price of silicon. All specifications are estimated in logarithms and with product-specific dummy variables. Absolute values of t-statistics are shown in parentheses below the parameter estimates. *** (**, *) denotes a 99% (95%, 90%) level of significance.

Table 7: Results for the Production Policy Equation

	Levels		First differences	
	IV-GLS	Lagged output	Arellano Bond	Bover
Variable	(1)	(2)	(3)	(4)
Constant	-1.6254	401.9645	-0.0934	-0.2876
	(-3.71)***	(6.97)***	(-0.40)	(-1.64)
Demand shock	-0.0314	0.0216	0.0157	0.0199
	(-3.77)***	(2.40)*	(0.85)	(1.40)
Log(Price of silicon)	0.3221	-0.3890	-0.4008	-0.3294
	(5.53)***	(-5.77)***	(-5.84)***	(-6.48)***
Log(Lagged number of firms)	3.3987	0.5112	0.0305	0.2460
	(48.66)***	(4.92)***	(0.18)	(2.02)*
Log(Cumulative past output)	0.8810	0.5403	2.2286	0.6147
	(66.81)***	(5.00)***	(3.55)***	(1.18)
Log(Cumulative past output of other firms)	-0.6363	-0.1692	-1.0640	0.2987
	(-45.72)***	(-4.83)***	(-2.00)*	(0.68)
AR(1)	0.8690		-0.2051	-0.0408
	(96.31)***		(-2.43)*	(-0.32)
Lagged output		0.5446		
		(7.81)***		
Time trend	-0.0000	-0.0730		
	(-2.89)**	(-6.93)***		
Dummy variable for 16K	-3.2743	-7.1958	-0.1134	0.1153
	(-17.34)***	(-7.03)***	(-0.47)	(0.64)
Dummy variable for 64K	-2.4125	-6.2205	-0.1856	0.1045
	(-12.96)***	(-7.13)***	(-0.74)	(0.56)
Dummy variable for 256K	-3.1350	-5.3182	-0.1364	0.1168
	(-16.89)***	(-7.03)***	(-0.56)	(0.65)
Dummy variable for 1MB	-2.8395	-4.4388	-0.1119	0.1415
	(-15.40)***	(-7.06)***	(-0.46)	(0.78)
Dummy variable for 4MB	-2.6404	-3.6212	-0.1293	0.1397
	(-14.43)***	(-7.12)***	(-0.53)	(0.77)
Dummy variable for 16MB	-2.2012	-2.8590	-0.0687	0.1630
	(-12.30)***	(-7.17)***	(-0.29)	(0.91)
Dummy variable for 64MB	-1.1914	-1.9923	-0.1345	0.0898
	(-6.74)***	(-6.92)***	(-0.56)	(0.50)
Dummy variable for 128MB	-0.4568	-1.4939	-0.2641	-0.0100
	(-2.57)*	(-6.35)***	(-1.08)	(-0.05)
Dummy variable for 256MB	0.4426	-0.8835	-0.1572	0.0918
	(2.48)*	(-4.54)***	(-0.64)	(0.49)
Number of observations	3635	3790	3463	3464
R-squared adjusted	0.89	0.89	0.01	0.17

Table 7 presents the instrumental variable estimation results for the incumbents' policy function in levels and in first differences. The dependent variable is the firm-specific output. Explanatory variables are the demand shock, the price of silicon, the lagged number of firms in the market, firm-specific past cumulative output and cumulative past output of all other firms. The estimations in the first two columns also has the firm-specific lagged dependent variable on the right hand side. All specifications are estimated in logarithms and with product-specific dummy variables. Absolute values of t-statistics are shown in parentheses below the parameter estimates. *** (**, *) denotes a 99% (95%, 90%) level of significance.

Table 8: Results for the Entry and Exit Equations

	Entry	Exit	Exit	Exit
Variable	(1)	(2)	(3)	(4)
Constant	-355.0836	-520.6153	-604.6940	-554.8208
	(-6.32)***	(-6.23)***	(-7.50)***	(-6.46)***
Demand shock	0.0285	-0.0728	-0.0718	-0.0864
	(0.74)	(-1.29)	(-1.34)	(-1.50)
Productivity shock (IV-GLS)		-0.2689		
		(-3.85)***		
Productivity shock (IV-L)			-0.2879	
			(-5.01)***	
Productivity shock (IV-FD)				-0.2121
				(-2.53)*
Log(Lagged industry output)	0.0448			
	(0.95)			
Log(Price of silicon)	0.2430	-0.3033	-0.3544	-0.3235
	(0.94)	(-1.06)	(-1.27)	(-1.11)
Log(Number of firms)	0.1534	0.1597	0.3201	0.2747
	(0.95)	(0.59)	(1.24)	(1.00)
Log(Cumulative past output)		-0.3392	-0.3053	-0.3314
		(-5.95)***	(-5.83)***	(-5.47)***
Log(Cumulative past output of other firms)		0.2313	0.0264	0.1564
		(2.24)*	(0.35)	(1.40)
Time trend	0.0605	0.0893	0.1052	0.0953
	(6.38)***	(6.28)***	(7.67)***	(6.52)***
Dummy variable for 16K	-1.1229	-1.1185	1.3730	-1.0262
	(-1.18)	(-3.52)***	(2.47)*	(-3.21)**
Dummy variable for 64K	-2.3335	-2.5804	-0.2745	-2.5720
	(-2.31)*	(-6.17)***	(-0.46)	(-6.07)***
Dummy variable for 256K	-2.8044	-4.5526	-2.4645	-4.5108
	(-2.72)**	(-6.45)***	(-3.14)**	(-6.27)***
Dummy variable for 1MB	-3.3529	-5.6680	-3.7131	-5.7214
	(-3.15)**	(-6.46)***	(-3.99)***	(-6.39)***
Dummy variable for 4MB	-4.1249	-6.2409	-4.3477	-6.3371
	(-3.61)***	(-6.64)***	(-4.43)***	(-6.62)***
Dummy variable for 16MB	-4.7035	-6.6688	-4.9653	-6.8141
	(-4.02)***	(-6.59)***	(-4.76)***	(-6.62)***
Dummy variable for 64MB	-5.0822	-6.7840	-5.3483	-7.0466
	(-4.19)***	(-6.38)***	(-4.88)***	(-6.48)***
Dummy variable for 128MB	-5.2008	-8.3458	-5.8952	-8.5164
	(-4.25)***	(-6.83)***	(-5.10)***	(-6.82)***
Dummy variable for 256MB	-5.2763	-8.0599	-6.8435	
	(-4.18)***	(-6.44)***	(-5.52)***	
Dummy variable for 512MB	-5.8450			
	(-4.29)***			
Dummy variable for 1GB	-4.8602			
	(-3.27)**			
Number of observations	2441	4321	4509	4086
Pseudo R-squared	0.23	0.28	0.28	0.27

Table 8 presents the estimation results from the probit models of the entry and exit distribution. In the entry model (columns 1), the dependent variable is an indicator variable, which is equal to one when a firm enters the market and zero before. Explanatory variables are the demand shock, price of silicon, number of firms, and a time trend. In the exit model (columns 2-4), the dependent variable is an indicator variable, which is equal to one when a firm exits the market and zero before. Explanatory variables are the demand shock, productivity shock, price of silicon, number of firms, cumulative past output, and cumulative past output of other firms. All specifications are estimated in logarithms and with firm-specific and product-specific dummy variables. Absolute values of t-statistics are shown in parentheses below the parameter estimates. *** (**, *) denotes a 99% (95%, 90%) level of significance.

Table 9: Results for the Structural Parameters in the (Marginal) Cost Function

Variable	(1)	(2)
Constant	θ_0	4.699
		(7.06)***
Economies of scale	θ_1	-1.032
		(215.52)***
Learning effects	θ_3	-0.0049
		(2.83)***
Spillovers	θ_4	-.00040
		(2.66)***
Number of observations		304

Table 9 presents the structural parameters in the (marginal) cost function. Absolute values of t-statistics are shown in parentheses below the parameter estimates. *** (**, *) denotes a 99% (95%, 90%) level of significance.

B Appendix: Figures

Figure 1: Average DRAM selling prices in USD, 1974-2004

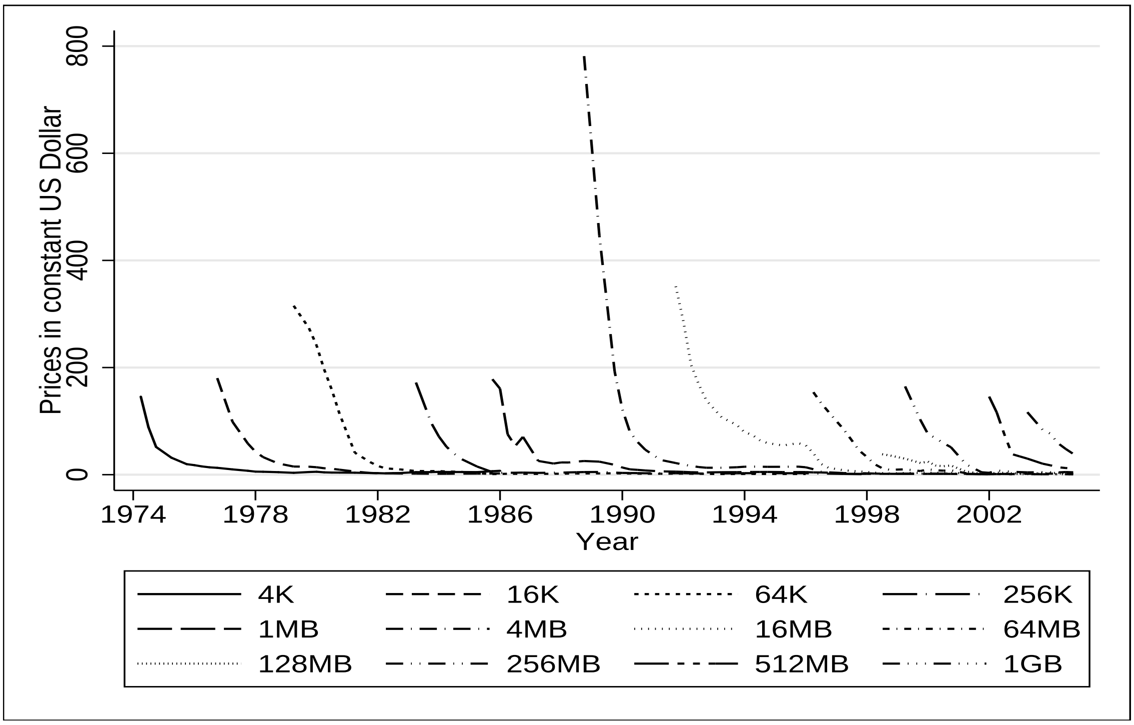


Figure 2: Industry units shipped, 1974-2004

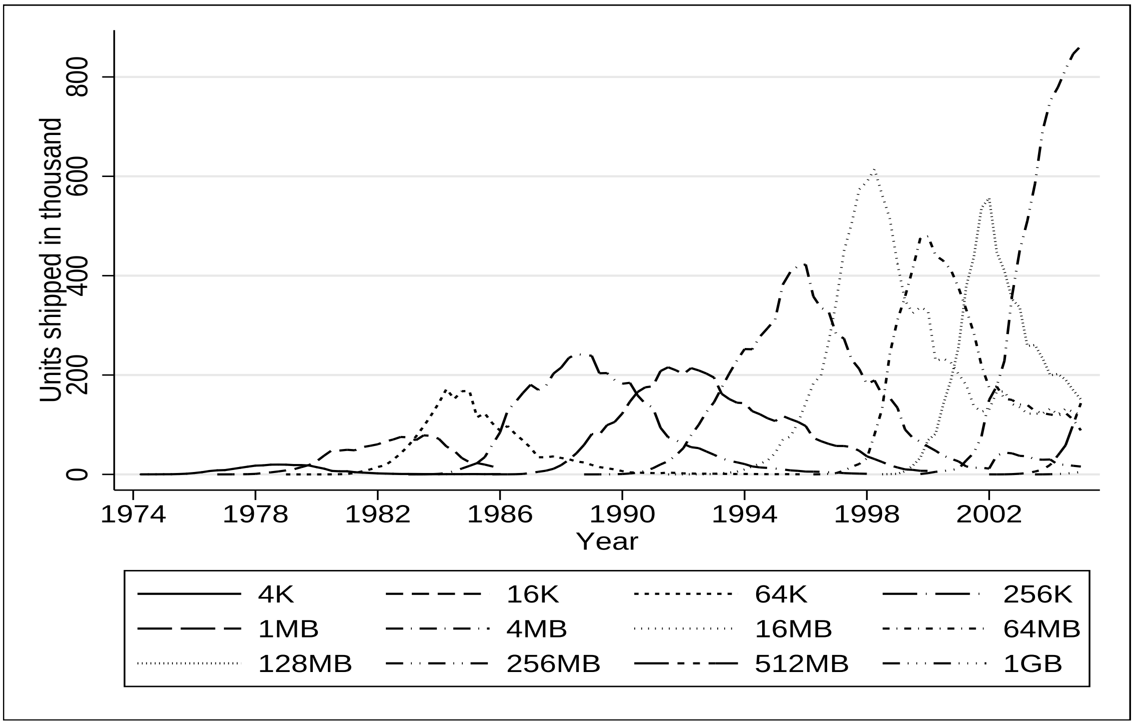


Figure 1b: DRAM selling prices in USD over different generations

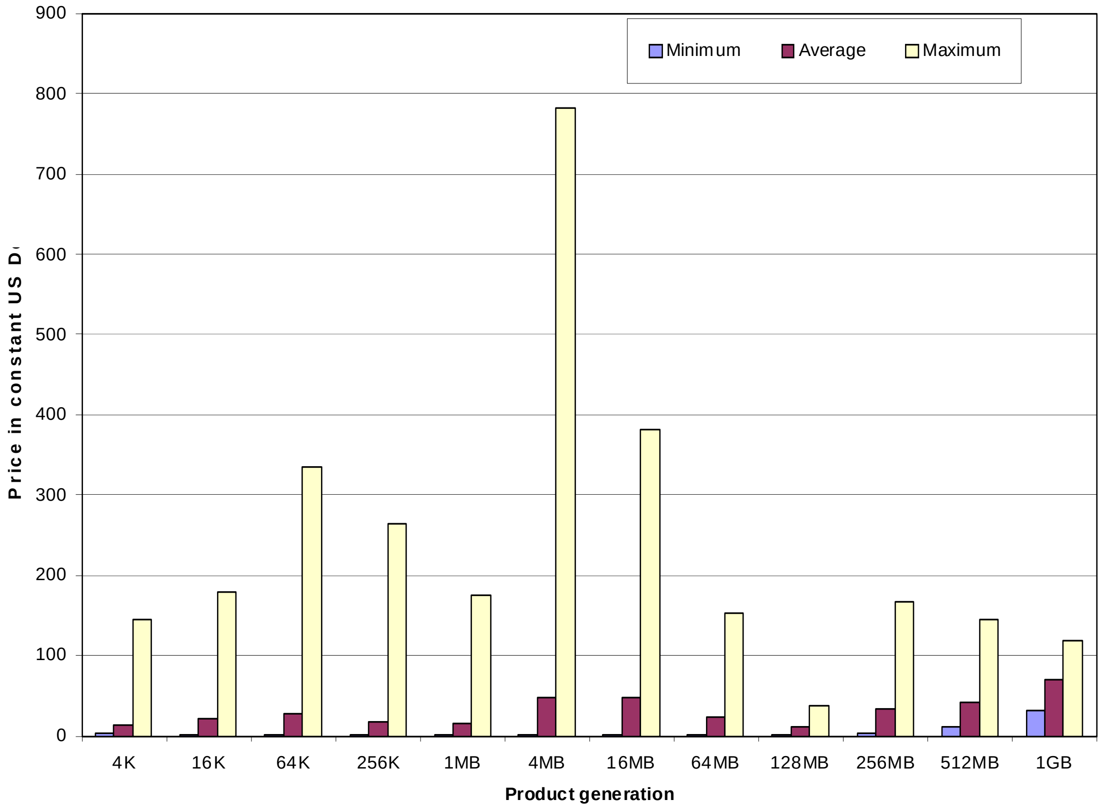
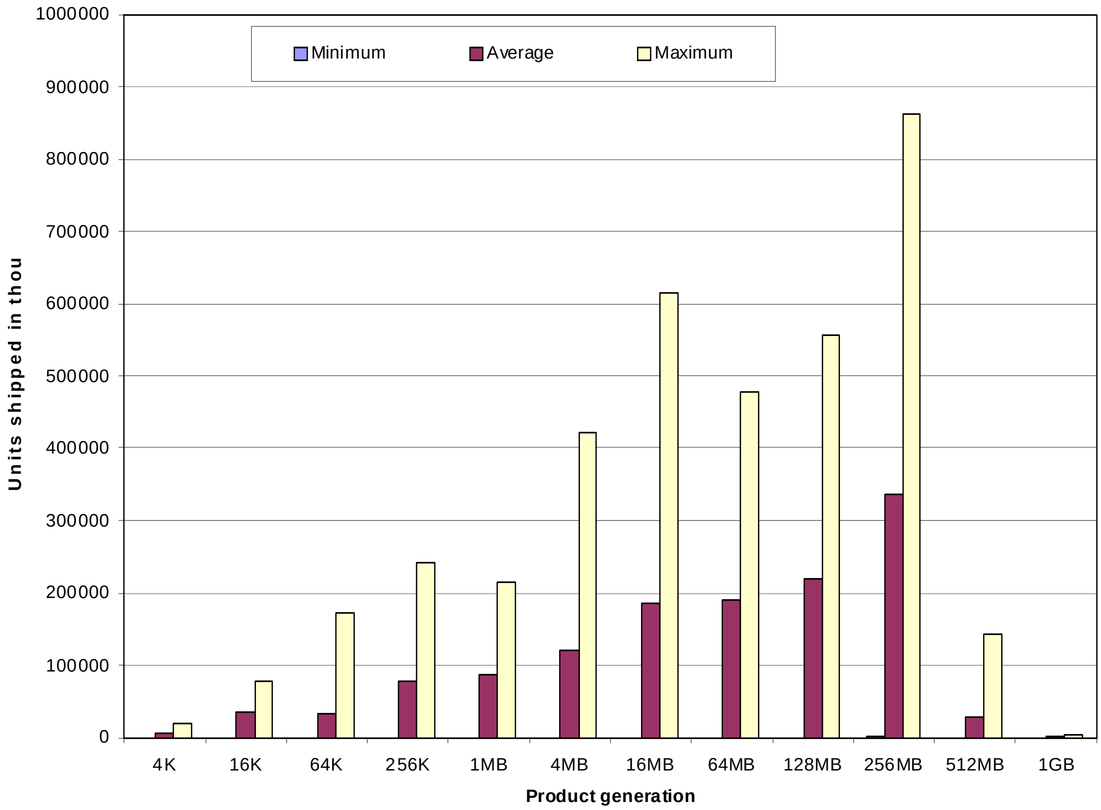


Figure 2b: Industry Units shipped in the DRAM industry over different generations



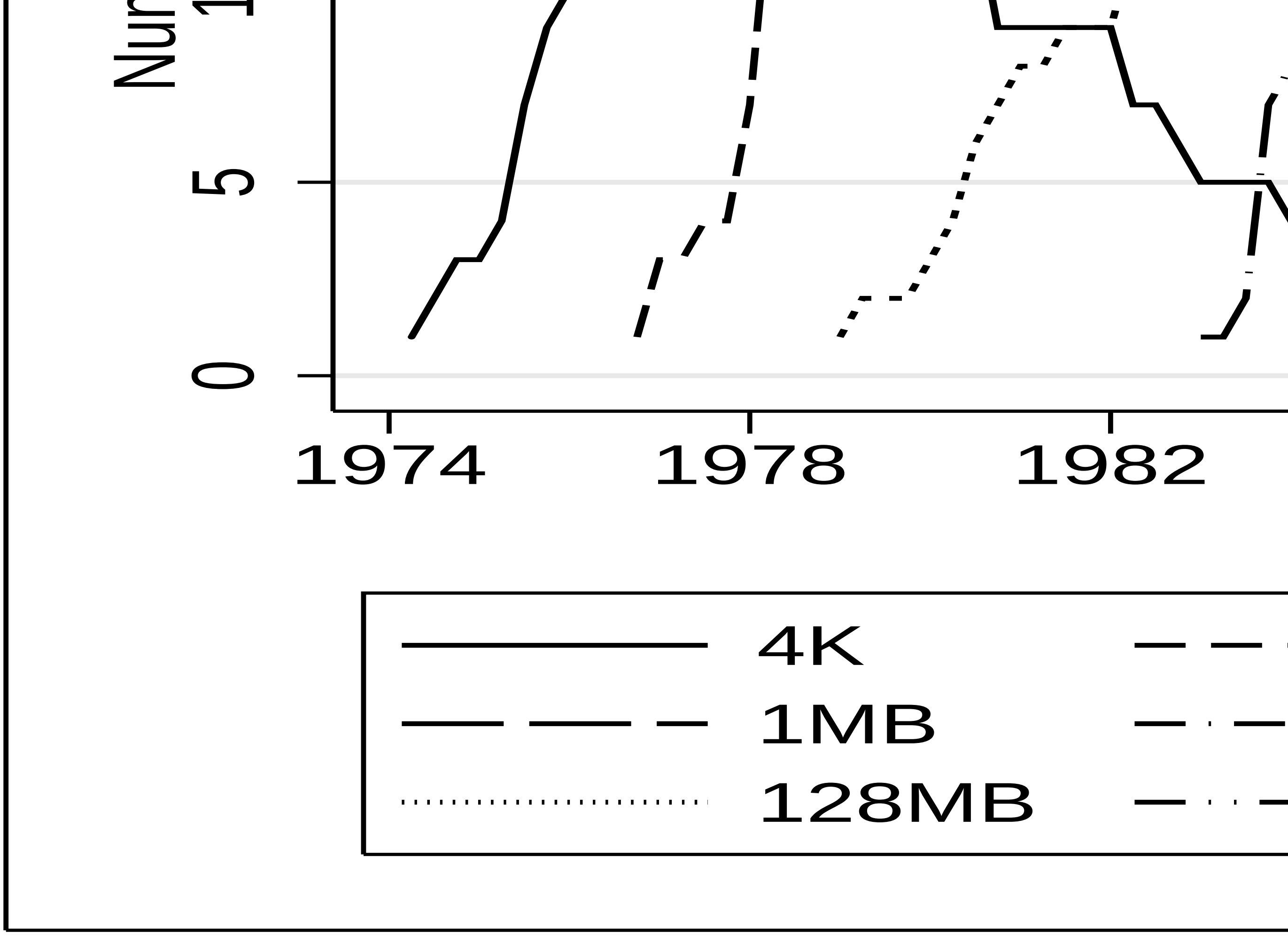

0
5
1974
1978
1982
4K
1MB
128MB

Figure 3b: Number of firms in the DRAM industry over different generations

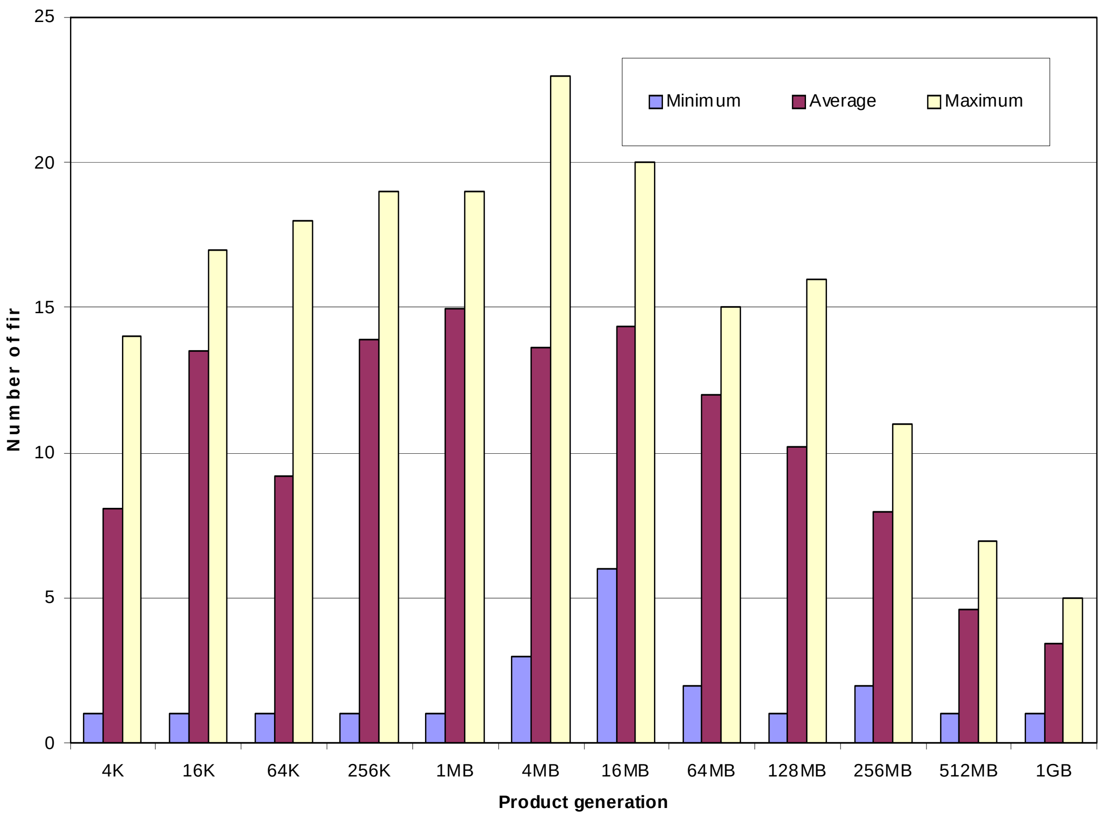


Figure 3c: Number of firms in the DRAM industry over different generations