AN EMPIRICAL EVALUATION OF THE 1986 SEMICONDUCTOR TRADE ARRANGEMENT

Sangin PARK*

Department of Economics
SUNY at Stony Brook
Stony Brook, NY 11794-4384
Tel: 631-632-7559
Fax: 631-632-7516
Email: sanpark@notes.cc.sunysb.edu
Web: http://ms.cc.sunysb.edu/~sanpark

First Draft: December 2000
This Draft: November 2002

This paper aims to empirically evaluate the extent to which the 1986 Semiconductor Trade Arrangement (STA) affected the Japanese producers' behavior and thus the evolution of the markets for the 256K and the 1M Dynamic Random Access Memory (DRAM) chips. We find that overall the STA facilitated the collusive behavior of the Japanese producers even beyond the myopic optimization: during the STA period, the Japanese firms' mark-ups were on average 2.3 (1.3) times as high as the myopic-optimization mark-ups and 6.1 (11.8) times as high as the learning-curve optimization mark-ups in the case of the 256K (1M) DRAM. In contrast, later-entering non-Japanese firms fully utilized the learning-curve optimization, rapidly expanding their shares in the worldwide (especially 1M DRAM) market. Our estimation results also suggest 'counter-product-cyclic mark-ups' consistent with learning-curve optimization behavior but do not support any spillover effects of learning by doing in this industry.

KEYWORDS: Semiconductor Trade Arrangement; price floor; learning by doing; product-cyclic competition; DRAM.

JEL Classification: F13; F12; L63.

* I am grateful for helpful comments from the participants of the 2001 North American Winter Meeting of the Econometric Society, the 2001 NBER Summer Institute International Trade and Investments Program, and the 2001 Econometric Society European Meeting. All errors and omissions are my own.

1. INTRODUCTION

The semiconductor industry, especially the Dynamic Random Access Memory (DRAM) chips industry, has drawn attentions from economists and policy makers for the last two decades. The attractions to this industry are mainly due to, among many reasons, the significance of learning-by-doing effects and the trade conflicts between the U.S. and Japan. The Semiconductor Trade Arrangement (STA) was one of the most comprehensive strategic trade agreements reached by the U.S. and Japan, which was implemented for 5 years from the forth quarter of 1986 (1986:4) to the third quarter of 1991 (1991:3). In this paper, we aim to empirically evaluate the extent to which the STA affected the Japanese producers' behavior and the evolution of the markets for the 256K DRAM and the 1M DRAM.

The organization of the paper is as follows. In section 2, we will first provide a brief review of the DRAM industry, some empirical facts and related theoretical literature. Then it will be shown that due to the learning-curve optimization and product-cyclic competition, the DRAM industry without any restrictions can be characterized by dynamic oligopoly. In section 3, we will briefly review the STA and its institutional restrictions which will be further utilized to develop the behavioral equations of Japanese and non-Japanese producers during the STA period. In section 4, we will report the estimation results of these behavioral equations. Based on these estimates, we can calculate Japanese and non-Japanese firms' actual mark-ups during the STA period as well as simulate the myopic-optimization mark-ups as reference values. The comparison of these mark-ups will reveal whether the trade restrictions of the STA simply constrained Japanese firms' learning-curve optimization or further facilitated Japanese firms' collusive behavior. In addition, we will be able to understand whether the typical movement of DRAM prices documented in section 2 is mainly motivated by the learning-curve optimization or by the product-cyclic competition. We will conclude the paper in section 5.

2. THE DRAM INDUSTRY: FACTS AND THEORIES

2.1. The DRAM Industry

DRAM is a particular kind of memory chips used in electronics products, mainly in computers and telecommunications equipments, in which the binary data can be stored and retrieved from any location. The manufacturing of DRAM includes repeated processes to create successive layers of insulating and conducting materials on a chemically treated silicon wafer by photographic techniques. The successfully processed wafer can then be diced into individual chips. The manufacturing process of DRAM is very sensitive: the mix of temperature, timing, chemicals and pressure must be precisely controlled, and a clean fabrication environment is necessary to avoid from dust, bacteria and unwanted contaminants. Improving 'yields' — the percentage of the functional chips on a processed silicon wafer — requires engineers' analyses and experiments on yield losses and a repetition of operations. Typically, yields can be very low, perhaps under 1 percent, in the early development stage and then continue to rise to as high as 90 percent (see Dohlman (1993)). So the higher the yield is, the lower the cost per chip. Hence, the cost per chip decreases with the cumulative volumes of chips produced (*learning-by-doing effects*).

Another important characteristic of the DRAM industry is the *product-cyclic competition*. Typically, each generation of DRAM is introduced by a leading firm, induces (sequential) entries of up to 20 producers, and then is completely replaced by the next generation of DRAM (see figures 1 and 2). In fact, as more firms enter over the product cycle, the market becomes less concentrated (see figure 3). The product cycle of a generation of DRAM, however, is *not* generated by the typical major innovation and imitations. The sequential entries of the DRAM producers are mainly caused by the stochastic nature of firms' succeeding in their own R&D in the case of early entries and then by the licensing of requisite production technology (called second sourcing) from the earlier innovators in the case of late entries (see Dick (1992); Irwin (1998)). Usually, the firms, which succeed in the innovation for a new-generation DRAM, come up with various different physical designs of the chip's internal structure. However, the new-generation DRAMs produced by different firms are not distinguishable in performance and nearly perfect substitutes to each other. Hence, DRAMs are considered homogenous commodities. A generation of DRAM, however, has a short economic life due to the rapid pace of technological change. Large investments in R&D rapidly improve the technology of manufacturing processes, which makes it possible to pack more circuit components onto a single chip. Intel introduced the first 1K DRAM in 1970, and since then DRAM capacities increased almost every three years by multiple of four: 4K in 1974, 16K in 1976, 64K in 1978, 256K in 1982, 1M in 1985, 4M in 1988, 16M in 1991, and so forth.

DRAM is usually sold through three channels: direct sales to larger users such as electronic equipment manufacturers, sales via authorized distributors, and the gray or spot market. As of 1989, direct sales count for almost 70 percent of DRAM sales, sales via authorized distributors 15 percent, and the gray market the rest of 15 percent of sales (see Flamm (1993a)). The direct sales and sales via authorized distributors are called 'contract sales' in which sellers commit, usually a quarter in advance, to supply certain quantity at some specified price. The major objective of these contract sales is the assurance of the delivery of committed quantities, especially during the period of shortage.[1] In the gray market, on the other hand, customers place orders and have products delivered immediately from suppliers such as independent distributors, brokers, and manufacturers. However, unlike contract sales, the quality of DRAM purchased in

[1] The contract prices are usually binding in the beginning of the duration of the contract, but for the rest of the period, purchasers usually can renegotiate the prices with suppliers for the changes of the market condition.

the gray market is not assured by the DRAM manufacturers.

2.2. Empirical Facts and the Literature

Typically, as illustrated in figure 4, the price of each generation of DRAM drops rapidly in the early stage of the product cycle, stays around at a lowest level, and then rises a little bit in the late stage of the product cycle. This typical price movement may be consistent with either learning-by-doing effects or product-cyclic competition. However, the typical price movement as well as those of entry and market concentration ratios (see figures 2 and 3) in the DRAM industry does not seem to be supported by the learning-curve-optimization behavior predicted in the theoretical literature. As well known, Spence (1981) showed a flat equilibrium price with learning-by-doing effects when the producer's discount factor is one. Empirically, an almost flat pricing pattern with learning by doing is documented in the market for wide-bodied commercial aircraft in Benkard (2000a). The previous theoretical studies have also predicted increasing market concentrations with strong learning effects. Dasgupta and Stiglitz (1988) showed that when learning is strong, a firm that possesses an initial cost advantage tends to increase its market share over time, and hence the industry will experience growing concentration, leading possibly to monopoly. Cabral and Riordan (1994), based on a price-setting duopoly of differentiated products with a single unit of demand, proved a weak increasing dominance. Moreover, Athey and Schmutzler (2001), allowing for downward-sloping demand and more than two firms, reached a similar conclusion when either the marginal learning effect or the discount factor is small enough. The typical pattern of price movement, entry and concentration ratio, therefore, seems to imply that learning-curve optimization is not strong, or at least product-cyclic competition dominates in the DRAM industry.

Apart from the traditional learning-curve-optimization model, Baldwin and Krugman (1988) and Baldwin (1994) proposed a new model to explain the price movement of DRAM with learning effects. In this new model, the DRAM producers are assumed to be pre-committed to their full-capacity productions with increasing yields. Hence rising yields of all firms mean increasing outputs over time, and thus as time passes, average costs fall at about the rate of output increase, but price falls at a rate less than that of output increase due to the elastic demand of DRAM. Baldwin and Krugman (1988) and Baldwin (1994) went further to argue that free entry will push down the DRAM producers' profits to the level of zero, and hence the average cost (not marginal cost) is below the price in the early stages but above the price in the later stages. Although the model developed in Baldwin and Krugman (1988) and Baldwin (1994) successfully predicts falling prices with learning effects, the assumption that firms cannot adjust their production levels is a restrictive one. In addition, despite increasing sales predicted by this model, total sales of each generation of DRAM typically rise and then fall. Moreover, the simulation study in Baldwin and Krugman (1986) turned out to predict too small a number of active firms in the 16K DRAM industry.

2.3. A Structural Model of Dynamic Oligopoly

The DRAM industry (without any restrictions) can be characterized by *dynamic* (rather than static) oligopoly since in the presence of learning-by-doing effects, a firm takes into account that its current production level will affect its future production costs and thus its future profits. In this subsection, we develop a structural model of dynamic oligopoly which is rich and flexible enough to incorporate the effects of both learning-curve optimization and product-cyclic competition into the equilibrium behavior of the DRAM producers. We begin by assuming that firms compete in their output levels. In other words, instead of assuming pre-committed full-capacity productions as in Baldwin and Krugman (1988) and Baldwin (1994), we consider quantity as the strategic variable in the dynamic oligopoly of the DRAM industry. There are several good reasons that quantity instead of price is considered as the strategic variable in the case of DRAM. As discussed above, the DRAM chips are considered homogeneous commodities, and most of them are sold by contracts which usually specify quantities and prices a quarter in advance. The major objective of these contract sales is for sellers to assure buyers of the committed quantities, and thus DRAM manufactures fix their production schedules a quarter in advance. In addition, the manufacturing of DRAM is clearly capacity constrained: there is usually a one-to-three-year time lag between the fabrication equipment purchase and the initial production (see Dohlman (1993)). In general, however, it is believed that capacity constraint has not been binding in the DRAM production. According to the report by Semiconductor Industry Association, the capacity utilization in wafer fabrication for all semiconductor products was 43 to 78 percent over 1978 – 1992.

In order to reflect learning-by-doing effects in the production process, we specify the marginal (or average variable) cost function of a firm as follows. Let q_{jt} denote firm j's output level, B_{jt} denote firm j's cumulative production, and ω_{jt} be the firm's cost characteristics in (discrete) time t. Note that $B_{jt+1} = B_{jt} + q_{jt}$. Let mc_{jt} denote firm j's marginal cost in time t. Then we assume that the marginal cost is a function of B_{jt} and ω_{jt} and is decreasing in B_{jt} at a decreasing rate.[2]

ASSUMPTION 1: $mc_{jt} = mc(B_{jt}, \omega_{jt})$ with $\partial mc_{it} / \partial B_{jt} < 0$ and $\partial^2 mc_{it} / \partial B_{jt}^2 > 0$.

Let p_t denote the market price of the homogenous products in time t and X_t denote the demand-side exogenous state variables. Then the inverse demand function is given as: $p_t = p(Q_t, X_t)$ where $Q_t = \sum_j q_{jt}$. Let Ω_t denote a vector of exogenous state variables, i.e., $\Omega_t = (X_t, \omega_t)$, where ω_t is a vector of all the firms' cost characteristics in time t. Assume that Ω_t evolves by following a

[2] We will consider industry-wide and inter-generational spillover effects of production experiences in section 4.

(stationary) Markov process.[3] Then under some regularity conditions,[4] a firm's optimal decision rule solves for the following Bellman equation:

$$(2.1) \quad V_j(\Omega_t, B_t) = sup_{q_{jt}} \{ (p(Q_t, X_t) - mc(B_{jt}, \omega_{jt}))q_{jt} + \phi E[V_j(\Omega_{t+1}, B_t + q_t) \mid \Omega_t]\},$$

where ϕ is the DRAM manufacturer's discount factor, and $E[\cdot|\cdot]$ is the conditional expectations operator.

With appropriate differentiability assumptions, we obtain the first-order conditions of the Bellman equation in (2.1), leading to the optimal price-cost margin as follows.

$$(2.2) \quad p_t - mc_{jt} = -\frac{\partial p(Q_t, X_t)}{\partial q_{jt}} q_{jt} - \phi E[\frac{\partial V_j(\Omega_{t+1}, B_{t+1})}{\partial B_{jt+1}} \mid \Omega_t].$$

Recall that $\partial B_{it+1} / \partial q_{jt} = 0$ if $i \neq j$ due to the usual Cournot-Nash assumption. Let $\Pi_{jt} = [p_t(X_t, \sum_i q_i(\Omega_t, B_t)) - mc(B_{jt}, \omega_{jt})]q_j(\Omega_t, B_t)$, where $q_j(\Omega_t, B_t)$ is the optimal policy of firm j. Then, the second term of the right-hand side of equation (2.2) can be rewritten as follows.[5]

---

[3] As will be detailed in section 4, Ω_t includes business-cyclic and product-cyclic shifts of demands for DRAMs. Generally, some demand shifters such as the aggregate quarterly GDP might have a trend of growth. However, a stationary Markov process assumption fits our case well since our data coverage is relatively short (5 years) and, moreover, the pattern of each generational DRAM prices and sales does not indicate any trend of growth.

[4] For the regularity conditions of (2.2) and the conditions for the existence of a Markov perfect equilibrium generated by dynamic optimization behavior based on (2.2), refer to Ericson and Pakes (1995) in a general case or Benkard (2000b) in the case of dynamic oligopoly.

[5] Note that $(\partial q_{jt+1} / \partial B_{jt+1})(\partial / \partial q_{jt+1})(\Pi_{jt+1} + \phi E[V_j(\Omega_{t+2}, B_{t+1} + q_{t+1}) \mid \Omega_{t+1}]) = 0$ since the (indirect) effects of a firm's production experiences on its own current and future costs (and thus profits) through the changes of its own output level will be fully adjusted and internalized in the firm's optimal quantity setting in the following time period (see, for example, the Benveniste and Scheinkman Theorem (Theorem 4.10) in Stocky, Lucas and Prescott (1989)).

(2.3) $$-\phi E[\frac{\partial V_j(\Omega_{t+1}, B_{t+1})}{\partial B_{jt+1}} \mid \Omega_t] = \phi E[\frac{\partial mc(B_{jt+1}, \omega_{jt+1})}{\partial B_{jt+1}} q_{jt+1} \mid \Omega_t]$$

$$-\phi E[\sum_{i \neq j} \{\frac{\partial \Pi_{jt+1}}{\partial q_{it+1}} + \phi \frac{\partial V_j(\Omega_{t+2}, B_{t+1}+q_{t+1})}{\partial q_{it+1}}\} \frac{\partial q_i(\Omega_{t+1}, B_{t+1})}{\partial B_{jt+1}} \mid \Omega_t].$$

Equation (2.2) indicates that the firm's margin in dynamic oligopoly can be decomposed into a *static* part (the first term of the right-hand side) and a *dynamic* part (the second term of the right-hand side). The dynamic part, as indicated in (2.3), further depends on the effects of increased production experiences on its own profits through the reduction of its own future marginal cost (*direct effect*) and through the changes of the other firms' future optimal output levels (*indirect effect*). Since $\partial mc_{jt+1} / \partial B_{jt+1} < 0$, the direct effect will have a negative sign. The indirect effect may have a negative or a positive sign. However, it would be exceptional that the dynamic part of the margin as a whole had a positive sign.[6] In the paper, we will call the static part of the margin *myopic-optimization margin* and the static part of the margin plus the direct effect of the dynamic part *learning-curve-optimization margin*. Note that myopic-optimization margin is the optimal margin when the firm behaves myopically while learning-curve-optimization margin can be derived from the Euler equation when the strategic interactions are negligible.

In the quantitative analyses of section 4, we will instead use a mark-up (i.e., a margin divided by the price) to measure the market power of the DRAM producers over the product cycle. As more entries occur over the product cycle, we may expect the static part of the mark-up to get smaller. On the other hand, as observed in the yield rates of the DRAM manufacturing, the effects of an increased experience will be greater in the earlier stages of the manufacturing, and thus the negative (direct) dynamic part of the margin will get smaller in an absolute value over

[6] Fudenberg and Tirole (1983) showed the possibility that the strategic incentives may induce firms to choose decreasing output paths and thus increasing price paths in the presence of learning by doing.

the product cycle. Moreover, if this decrease in the (negative) dynamic part of the margin (in an absolute value) is greater than the decrease in the price over the product cycle, the (negative) dynamic part of the margin will get smaller in an absolute value over the product cycle. Hence, depending on the rates of the declines of the positive static part of the mark-up and the negative dynamic part, we may observe different patterns of the mark-up over the product cycle.[7] Hence, to quantify the importance of the learning-curve optimization and the product-cyclic competition in the DRAM producer's optimal quantity-setting behavior, we have to obtain consistent estimates of marginal production costs. Due to the indirect dynamic part of the mark-up of (2.3), however, we cannot generally obtain a feasible estimation procedure in dynamic oligopoly (see Park (forthcoming); Berry and Pakes (2001)).[8] Yet, in our specific case, the restriction imposed by the STA will be utilized to obtain a feasible estimation procedure for the marginal production costs.

3. A MODEL FOR THE SEMICONDUCTOR TRADE ARRANGEMENT

3.1. The 1986 Semiconductor Trade Arrangement

In the 1980s, the semiconductor industry became a subject of the trade conflicts between the U.S. and Japan. Japanese firms gained a 40 percent share of the 16K U.S. DRAM market

[7] Previous empirical studies of the DRAM manufacturing have implicitly restricted the DRAM producer's price-cost margins in the specification of models. Often prices have been employed as a proxy for marginal costs (for example, Gruber (1992); Dick (1991)). The underlying assumption of this approach is that margins are constant both across firms and over time. In the case of DRAM, however, we expect that learning-curve optimization as well as increasing intensities of competition must affect firms' margins over the product cycle. Irwin and Klenow (1994) made a difference, considering different margins across firms and over time. However, their specification assumes that a firm's margins increases with its market share, which implies 'product-cyclic mark-ups': higher mark-ups in the earlier product cycle of DRAM and then lower mark-ups in the later product cycle. Hence in their specification, learning-curve optimization is not fully reflected.

[8] For the estimation of a structural model of dynamic oligopoly, Park (forthcoming) requires the two sets of instrumental variables satisfying an orthogonality condition in the filtering process while Berry and Pakes (2001) excludes an existence of unobservable cost-side shocks.

during the late 1970s and increased their share of the 64K DRAM market up to 70 percent in the early 1980s. The continuing ascent of Japanese world market shares and a cyclical decline of demand drove the U.S. semiconductor industry into a crisis in 1985-1986. The U.S. semiconductor producers accused the Japanese chip makers of dumping EPROM (Erasable Programmable Read-Only Memory) and DRAM chips in the U.S. and the third-country markets and complained to the U.S. government about Japanese barriers to imports of U.S.-origin semiconductors.[9] In addition, the U.S. government became increasingly worried about the dependency on Japan for the supply of DRAMs which were considered strategic inputs for many high-tech industries. These escalating tensions led the U.S. and Japanese governments to reach a strategic trade regime, called the STA, in July 1986, which was implemented for 5 years, from 1986:4 to 1991:3.[10] Under the arrangement, both governments agreed: (i) to encourage increased sales of foreign-origin semiconductors in the Japanese market;[11] (ii) to refrain from policies or programs which stimulated inordinate increases in semiconductor production capacity; and (iii) for the U.S. Department of Commerce (DOC) to administer a system of firm-specific price floors (also called Foreign Market Values) on Japanese DRAM and EPROM exports. For detailed discussions of the STA, the trade politics leading to the STA, and its aftermath, refer to Irwin (1996, 1998), Flamm (1993c, 1996), and Dohlman (1993).

It has been widely reported that among the three agreements, the system of firm-specific price floors was most effectively enforced. During the STA period, the system of firm-specific price floors for Japanese DRAM exports was imposed on the three generations of DRAM chips,

[9] Dick (1991), based on the data for the 64K DRAM in 1985:1 and the 256K DRAM in 1985:2, argued that Japanese firms were formally guilty of dumping under the U.S. trade law but made their output decisions based on learning-curve optimization rather than engaging in predation against the U.S. competitors. On the other hand, based on certain regularities in the history of the DRAM industry, Flamn (1993b) was more suspicious of the predatory dumping while Irwin (1998) was more inclined to a cyclical dumping.

[10] In 1991, the STA was extended for another five years. Instead of firm-specific price floors, the second STA set a new fast-track procedure supervised by Japanese government. In the new procedure, “Japanese companies were required to collect cost and price data and have it available to the U.S. and Japanese governments on fourteen days’ notice in the event of an antidumping case” (Flamm (1996), p.162).

[11] It is believed that there were “secret” side letters aiming that the market share of foreign-capital-affiliated companies in Japan would increase to at least 20 percent over the period of the STA (Flamm (1996)).

such as the 64K, the 256K and the 1M DRAMs. However, the focus was on the 256K and the 1M DRAMs. In mid 1980s, the 64K DRAM was already an old vintage product, while the 256K DRAM was in the mature peak stage, and the 1M DRAM in the early expanding stage of the product cycle. Although the firm-specific price floors were finally determined by the DOC, it was Japanese government which bore the burden of enforcing these price floors. As widely documented as in Irwin (1996, 1998), Flamm (1993c, 1996), and Dohlman (1993), the Ministry of International Trade and Industry (MITI) of Japan set *firm-specific quotas* for the Japanese productions of the 256K and the 1M DRAMs to implement these price floors. The MITI established the Supply and Demand Forecast Commission in September 1986, which, until it was disbanded in mid 1989, played a key role, not only publishing forecasts but also effectively determining the production levels of Japanese firms. Without much delay, the MITI began to effectively monitor to prevent exports at prices less than those price floors. The MITI tightened export licenses by January 1987 and increased efforts to eliminate gray market sales by February 1987. By the first half of 1987, the MITI successfully pressed Japanese firms to meet minimum export price guidelines even for third-country markets.

The impacts of price floors and corresponding quotas set for the Japanese firms during the STA period can be detected in the movement of the prices of the 256K and the 1M DRAMs in the worldwide market. As of 1986:4, the Japanese firms occupied 80.5 percent of the 265K DRAM sales and 88.7 percent of the 1M DRAM sales. Naturally, the impact of the STA was immediate and significant. The prices of the 1M DRAM and the 256K DRAM were immediately raised and hanged at an unusually high level during the STA period (see figure 4).[12] These boosts of prices are truly an exception for the typical price movement discussed in section 2. The comparison of the prices of different generational DRAMs over the same period of the product cycle clearly indicates these uncharacteristic movements of prices of the 256K and the 1 M

[12] Figure 4 also indicates an uncharacteristic price hike of the 1M DRAM in 1986:3, the quarter just before the STA was implemented.

DRAMs during the STA period. The 256K DRAM and the 1M DRAM experienced the STA from the 17th to the 36th quarter and the 6th to the 25th quarter of the product cycle, respectively. As shown in figure 5, from the 6th to the 16th quarter of the product cycle, in which the 1M DRAM was under the influence of the STA but the 256K DRAM was not, the 256K DRAM had experienced sharp price drops while the price of the 1M DRAM hanged on higher and declined slowly. From the 17th to the 36th quarter of the product cycle, in which the 256K DRAM was under the influence of the STA but the 64K DRAM was not, the 64K DRAM experienced sharp price drops and then began to rise in the late stage of the product cycle while the price of the 256K hanged on higher and declined slowly.

The impacts of these price floors and quotas can also be observed in the Japanese firms' market shares as well as the market concentration ratios (see figure 6). In the 17th quarter of the product cycle, in which the STA became effective for the 256K DRAM but was already in action for 10 quarters in the 1M DRAM case, the Japanese firms' share in the worldwide market was 80.5 percent in the 256K DRAM case but just 70.5 percent in the 1M DRAM case. Moreover, at the end of the STA, in which the 256K DRAM and the 1M DRAM were in the 36th quarter and the 25th quarter of the product cycle, respectively, this Japanese firms' share was reduced to be as low as 44.1 percent in the 256K DRAM case and 43.7 percent in the 1M DRAM case. Despite rapid decreases of Japanese market shares, especially of the 1M DRAM, during the STA period, the concentration ratio within the Japanese producers remained unusually high.[13] From the 6th to the 16th quarter of the product cycle, the concentration ratios of the worldwide market for the 1M DRAM was substantially lower than those of the 256K DRAM market. However, during the same period, within the Japanese firms, the concentration ratios of the 1M DRAM production were substantially higher than those of the 256K DRAM production although there were a little bit greater or equal number of Japanese producers in the manufacturing of the 1M DRAM (6 to

[13] In addition, a leading firm's share within the Japanese producers was substantially higher in the 1M DRAM production.

10 Japanese producers in the 1M DRAM production while 6 to 9 Japanese producers in the 256K DRAM production).

To sum up, the patterns of prices, Japanese market shares and concentration ratios during the STA period suggest that Japanese productions might have been substantially reduced to generate unusually high levels of the market price. In what follows, we will model the (restrained) behavior of Japanese producers applied to the data of the worldwide market during the STA period. Since firm-specific quotas set for Japanese DRAM productions might have worked as implicit VERs, it is questionable whether those quotas also constrained Japanese firms' domestic sales. As well documented in Flamm (1996, pp.242-254), in 1987 and 1988, there were significant regional price differentials in the 256K and the 1M DRAMs, especially between the Japanese market and the U.S. market although these regional price differentials disappeared especially from the second half of 1989. These regional price differentials may suggest different behavior of Japanese firms in the apparently unregulated domestic market in the earlier period of the STA. Hence, it will be ideal to apply the model developed below to the data for the non-Japanese (or the U.S.) market especially in the earlier period of the STA. However, to our knowledge, the firm-level sales data in the regional markets are not available. The possible biases caused by the lower Japanese market prices in 1987 and 1988 will be discussed in our estimation results of section 4.

3.2. Behavioral Assumptions during the STA Period

In the paper, we are mainly interested in the impacts of the STA on Japanese firms' behavior and thus the evolution of the markets for the 256K DRAM and the 1M DRAM. Specifically, we aim to analyze whether the price floors and quotas set during the STA period constrained Japanese firms' learning-curve optimization or further facilitated the collusive behavior.[14] During the STA period, the DOC set firm-specific price floors in each quarter, based on the projected (static) costs voluntarily reported by Japanese producers. It is believed that both the DOC and Japanese firms had iterated on to some set of procedures to generate firm-specific price floors although we do not know the exact procedure(s). Furthermore, to our knowledge, the information of these reported costs and the price floors set for each Japanese firm by the DOC during the STA period is still confidential.[15] Consequently, we cannot confirm whether the price floors were binding or actual prices were higher than these price floors.[16] However, even if this information is available, we do not know whether the reported costs are the *actually* projected ones. During the STA period, a Japanese firm might have had an incentive to report its cost to be lower than the actually projected one in order to explore learning-curve optimization. In this case, the price floor might have been binding. On the other hand, the price floors and quotas might have facilitated the collusive behavior of the Japanese firms. In the case, the price might have been higher than the price floor although the firm reported actually projected costs.

To empirically estimate these two incentives faced by the Japanese firms during the STA period, we will posit the following behavioral assumption on Japanese producers. Let p_t^e be the quarter-*t* price projected in quarter *t*-1, and mc_{jt}^e be Japanese firm *j*'s quarter-*t* (static) marginal cost projected in quarter *t*-1. In quarter *t*-1, the Japanese firm effectively (but in a restrictive manner) sets its price-cost margin of quarter *t*, say $\mu_{jt} = p_t^e - mc_{jt}^e$, by choosing a price-to-marginal-cost (P-MC) ratio, say λ_{jt} $(= p_t^e / mc_{jt}^e)$. Then, in quarter *t*, the regulator (the MITI) set the corresponding quotas to each Japanese firm, which credibly implemented the price-cost

[14] As discussed in Krishna (1989), trade restrictions may work as facilitating practices.

[15] Japanese firms were required to file in each quarter the public reports which contained, for confidentiality, 'ranged' estimates of various cost concepts, calculated by the value (reported to the DOC) plus or minus an error of up to 20 percent.

[16] It is widely speculated that the price floors were generally set quite close to the projected (average) costs submitted by Japanese firms. Based on this speculation and the ranged estimates of reported costs available to the public, Flamm (1996) found that the average price floors were overall lower than the Dataquest's worldwide average sales prices.

margins (and thus the price floors) set in quarter t-1. For example, suppose that p_t^e is a common knowledge in period t-1. Then, if the Japanese firm knows that the DOC will set a certain firm-specific price equivalent to a certain P-MC ratio, then it could effectively set this P-MC ratio by adjusting its reported marginal cost. Hence, we assume the price-cost margin of a Japanese firm during the STA period as follows.

$$(3.1) \qquad p_t - mc_{jt} = (\lambda_{jt} - 1)mc_{jt}^e .$$

As a focal case, we will consider that the interactions between the Japanese firm and the DOC in setting a firm-specific price floor might have constrained a Japanese firm to adopt a simple P-MC ratio such that the P-MC ratio, λ_{jt}, is an increasing function of a firm's cumulative production. Note that the cumulative production is the single most important factor in a firm's marginal cost (see the estimation results in section 4) and can be fairly well predicted (by the firm) in the previous quarter. Then, in our focal case, a Japanese firm's P-MC ratio in quarter t is specified as follows.

$$(3.2) \qquad \lambda_{jt} = \lambda_t(\lambda_0 + f(B_{jt})) ,$$

where λ_0 is a constant, λ_t is a time-specific positive constant, and f is an increasing real function. The time-specific constant λ_t reflects that the firm-specific P-MC ratio might also be affected by the phase of product cycle in time t. In the actual estimations of section 4, we will use a logarithm specification of f in (3.2) as a baseline case (and will also consider linear and exponential functional forms).

Note that the simple specification of the Japanese firm's P-MC ratio in (3.2) is consistent with the Japanese firms' incentives for learning-curve optimization and collusive behavior. First, a Japanese firm's incentive to explore learning-curve optimization is greater if its cumulative production experience is smaller. Second, a more efficient firm (i.e., a firm with more production experiences) would prefer a higher price-cost margin. Lastly, the simple formula for the P-MC ratio might go well with the trade restrictions as facilitating practices. In section 4, as alternative specifications, we will consider that the Japanese firm might have been able to commit to a more sophisticated P-MC ratio to set the price-cost margin closer to that of the myopic optimization or the learning-curve optimization.

Although no non-Japanese firm was restrained during the STA period, a single non-Japanese firm's share in the markets for the 256K and the 1M DRAMs was still limited.[17] Table 1 shows that non-Japanese firms' average share in the 1M DRAM market was between 0.4 percent and 6.3 percent, and the largest non-Japanese firm's share was between 0.4 percent and 15.8 percent. In the 256K DRAM case, these numbers were between 2.4 percent and 6.3 percent and between 7.5 percent and 16.5 percent, respectively. Hence, due to a small market share of any non-Japanese firm, we will assume that a small increase in the non-Japanese firm's production experience had negligible effects on the changes of the other firms' optimal output levels. In other words, we assume that the indirect effect of the dynamic part of the mark-up in (2.3) was negligible for any non-Japanese firm during the STA period. Note that non-Japanese firms were still considered to choose their current output levels strategically as in a usual (static) Cournot case. Then, the optimal price-cost margin of non-Japanese firm *i* during the STA period is the learning-curve-optimization margin as discussed in section 2.3.

$$(3.3) \qquad p_t - mc_{it} = -\frac{\partial p_t}{\partial q_{it}} q_{jt} + \phi E[\frac{\partial mc_{it+1}}{\partial B_{it+1}} q_{it+1} \mid \Omega_t].$$

[17] On the other hand, an individual non-Japanese firm's share in the markets for the 64K DRAM was substantial: the non-Japanese firms' average share in the 64K DRAM market was between 3.4 percent and 20.6 percent during the STA period, and the largest non-Japanese firm's share as high as 47.3 percent. As indicated in figures 2 and 3, there were substantial exits and increases of the market concentration in the 64K DRAM market during the STA period, which is a typical phenomenon for an old vintage DRAM.

Let $u_{it}\,(=\phi \frac{\partial mc_{it+1}}{\partial B_{it+1}} q_{it+1} - \phi E[\frac{\partial mc_{it+1}}{\partial B_{it+1}} q_{it+1} \mid \Omega_t])$ denote the projection error. Then by definition, $E[u_{it} \mid \Omega_t] = 0$.

3.3. Estimation Procedure

To estimate Japanese firms' P-MC ratios and the cost-side parameters of the 256K DRAM and the 1M DRAM manufacturing, we need to specify the marginal cost function and the demand function. We begin by assuming a marginal (or average variable) cost function with a constant learning elasticity as follows.[18]

$$(3.4) \qquad mc_{jt} = c_j B_{jt}{}^{\gamma} e^{W_{jt}\beta} + \varepsilon_{jt} \, ,$$

where γ (< 0) is called learning elasticity, c_j denotes an initial (firm-specific) inefficiency level, β is a vector of parameters, W_{jt} is the firm's observable cost characteristics, and ε_{jt} denotes a (unobservable) productivity shock in time t.[19] Note that the firm's exogenous cost characteristics, ω_{jt}, have two components, W_{jt} and ε_{jt}. Note also that c_j determines (conditioned on W_{jt} and ε_{jt}) the firm's marginal cost with no production experience. In section 4, we will test whether later-entering firms might have been more efficient initially (due to better equipment etc.). We also

[18] Hatch and Reichelstein (1997) discussed different specifications of the learning curve in the DRAM manufacturing. In actual estimations, we have tried different specifications of the marginal cost function such as a linear form, but we have found that this conventional specification fits better.

[19] In our specification, the marginal cost is additive in the productivity shock ε_{jt} instead of multiplicative in $\ln(\varepsilon_{jt})$. Either specification is basically for convenience in the estimation procedure. Typically, the multiplicative specification implies that the productivity shock is a constant percentage of marginal cost and thus decreases as the cumulative production increases. Since the identical distribution of ε_{jt} is not required in our GMM estimation, even this implication of the multiplicative specification can be consistent with our additive specification. However, our estimation results did not support the implications of the

assume that a Japanese firm's quarter-*t* marginal cost projected in quarter *t*-1, mc^{e}_{jt}, is the mean level of its marginal cost in quarter *t*.

ASSUMPTION 2: $mc^{e}_{jt} = c_j B_{jt}{}^{\gamma} e^{W_{jt}\beta}$.

The specification of the demand function for DRAM, however, is not straightforward since, as discussed in section 2, most of the DRAM sales are contract sales. Following previous studies such as Baldwin and Krugman (1988) and Irwin and Klenow (1994), however, we will simply assume a constant elasticity demand function for DRAM. Refer to section 4.1 for more discussions on this specification. Let η denote the elasticity of demand. Then the static part of the margin in (2.2) or (3.3) will be: $-(\partial p_t / \partial q_{jt}) q_{jt} = (S_{jt} / \eta) p_t$, where S_{jt} is the market share of firm *j* in time *t*.

Hereafter, let *j* denote a Japanese firm and *i* denote a non-Japanese firm. From this constant-elasticity demand function, the specifications of the P-MC ratio in (3.2), the marginal cost function in (3.4), and Assumption 2, we can rewrite the Japanese and the non-Japanese firms' price-cost margins in (3.1) and (3.3), generating the following estimating equations.

(3.5) $$\nu_{jt}(\theta_0) = p_t - \lambda_t(\lambda_0 + f(B_{jt}))c_j B^{\gamma}_{jt} e^{W_{jt}\beta}, \text{ and}$$

$$\nu_{it}(\theta_0) = (1 - \frac{S_{it}}{\eta}) p_t - \phi\gamma c_i B^{\gamma-1}_{it+1} e^{W_{it+1}\beta} q_{it+1} - c_i B^{\gamma}_{it} e^{W_{it}\beta},$$

where $\nu_{jt}(\theta_0) = \varepsilon_{jt}$, $\nu_{it}(\theta_0) = \varepsilon_{it} + u_{it}$, and θ_0 is the parameters we want to estimate including $c_j, \gamma, \beta, \lambda_0$ and λ_t.

multiplicative specification. We also tried '$c + \varepsilon_{jt}$' instead of ε_{jt} as an error term and found that c can be set

We will assume that the firm's productivity shock is mean independent of the observed product and cost characteristics.

ASSUMPTION 3: $E[\varepsilon_{jt} \mid X_t, W_t] = E[\varepsilon_{it} \mid X_t, W_t] = 0$ for all j, i and t.

Note that productivity shocks may be serially correlated and be correlated with cumulative outputs, B_{jt}. Since u_{it} is the projection error, we have: $E[u_{it} \mid \Omega_t] = 0$. Hence, we have: $E[u_{it} \mid X_t, W_t] = E[E[u_{it} \mid \Omega_t] \mid X_t, W_t] = 0$. This conditional moment restriction, coupled with Assumption 3, implies the following conditional moment restriction.

$$(3.6) \qquad E[\begin{pmatrix} \nu_{jt}(\theta_0) \\ \nu_{it}(\theta_0) \end{pmatrix} \mid X_t, W_t] = 0 .$$

Let $Z_t = (X_t, W_t)'$. Then the conditional moment restriction of (3.6) implies that the error terms are uncorrelated with any function, say H_{ji}, of Z_t. Define

$$(3.7) \qquad G_{J+I}(\theta) = \frac{1}{J+I}\sum_{j,i} m_{ji}(\theta) \text{ with } m_{ji}(\theta) = H_{ji}(Z_t)\begin{pmatrix} \nu_{jt}(\theta) \\ \nu_{it}(\theta) \end{pmatrix},$$

where J and I are the number of Japanese and non-Japanese observations in our data, respectively. Note that the conditional moment restriction in (3.7) implies that $G(\theta) \equiv E[m_{ji}(\theta)] = 0$ at $\theta = \theta_0$. Let $||y||_A = y'Ay$ for any vector y and a conformable matrix A. Then a Generalized Method of Moments (GMM) estimator, say $\hat{\theta}$, minimizes: $\| G_{J+I}(\theta) \|_{V_{J+I}}$, where V_{J+I} is a weighting matrix, converging to V in probability with $V^{-1} = E[m_{ji}(\theta_0)m_{ji}(\theta_0)']$. As well known,

to be zero based on the hypothesis testing statistics detailed in section 4.1.

under some regularity conditions in Hansen (1982) or Newey and McFadden (1994), the GMM estimator, $\hat{\theta}$, is consistent and asymptotically normal with the covariance matrix given by $\Lambda = (J + I)^{-1}(\Gamma' V\Gamma)^{-1}$, where $\Gamma = (\partial / \partial\theta')G(\theta_0)$.

4. EMPIRICAL FINDINGS

4.1. Data and Specification Issues

The data employed in this paper are quarterly observations, spanning 1986:4 to 1991:3. The data for prices and firm-level sales of the 256K and the 1M DRAMs are obtained from Dataquest, a private consulting firm. Prices are "average selling prices" (ASPs), the Dataquest estimates of the average worldwide billing prices in each quarter, while firm-level sales data are based on the factory revenue shipments in 1,000 units. The ASP is generally considered a good measure for the market price of DRAM. Refer to Flamm (1993a) for a detailed discussion of the properties of the ASP. The price data are deflated by the quarterly CPI obtained from the *Monthly Review of the U.S. Bureau of Labor Statistics*. The base quarter of these CPI deflators is 1967:1.[20] We also employ exchange rates as an observed cost shifter (observed cost characteristic, W_t). These data are obtained from the *International Financial Statistics*, IMF, and the *Monthly Bulletin of Statistics*, Taiwan. As demand shifters (observed product characteristics, X_t), we make use of the aggregate GDP of the OECD countries, the quarterly changes of this aggregate GDP, the vintage of the generation of a DRAM, the vintage squared, the vintage cubed, the vintage to the power four, and the quarterly dummy variables. The OECD countries' GDPs are obtained from *Quarterly National Accounts*, OECD. The aggregate GDP of the OECD countries and the

[20] We were not able to obtain any consistent series of quarterly CPIs whose base quarter is closer to the STA period, 1986:4 – 1991:3.

quarterly changes of this aggregate GDP are used to reflect the business-cyclic change of the demands for DRAM while the forth-degree polynomial series of vintage variables are used to reflect a bell-shaped demand pattern of each generation of DRAM over the product cycle.[21] Table 2 summarizes the demand and cost shifters in our estimations.

As discussed in section 3, we assume a constant elasticity demand function. Following previous studies such as Baldwin and Krugman (1988) and Irwin and Klenow (1994), we consider 1.8 as a baseline value of the elasticity of demand, η. The estimation results reported in table 4 are based on this baseline value case. To see if the conventional wisdom of '$\eta = 1.8$' is reasonable, we have checked the sensitivity of the estimation results to the value of η, ranging from 1.5 to 2.1. We have found our results very robust to these variations. Furthermore, we have specified the functional form of demand as: $log(Q_t) = \eta log(p_t) + X_t\rho + \xi_t$, where ξ_t is an unobserved demand-side shock, finding that the difference of $log(Q_t)$ - $\eta log(p_t)$ with $\eta = 1.8$ was very well explained by the demand shifters, X_t. The R^2 values of the OLS regression of $log(Q_t)$ – $1.8 \times log(p_t)$ onto the space spanned by X_t are 0.97 in the 1M DRAM case and 0.98 in the 256K DRAM case. Based on this specification of the demand function, we have also simultaneously estimated the demand function and the producer-side equations in (3.5). The estimates (1.83 in the 256K DRAM case and 1.8 in 1M DRAM case), however, turned out to have very big standard errors. This may be because the cost shifter (exchange rates) is a weak instrumental variable for the price in the demand function as will be implied in table 4.

In our baseline estimation, we also control the producer's discount factor, ϕ. Due to a short product cycle of DRAM, we will set $\phi = 0.95$ in our estimation. Note that ϕ can be estimated in our estimating equation of non-Japanese firms in (3.5). Indeed, ϕ is estimated to be 0.98 in the case of 256K DRAM and 0.97 in the case of 1M DRAM. However, the estimation

[21] In this paper, we do not explicitly consider the demand substitutions between a new generation DRAM and an old generation. In the early and the late stages of the product cycle, DRAMs usually serve for niche

results reported in table 4 remained almost unchanged in these estimations. Due to our shorter data covering different stages of the product cycle of 256K DRAM and 1M DRAM, however, we fix ϕ to be 0.95 for both generations of DRAMs in our focal case.

Based on our structural estimation of equation (3.7), we can apply the minimum distance statistic, say *D*, and the Wald statistic, say *W*, in Newey and McFadden (1994) to conduct several specification tests. Let the restriction under the null hypothesis be $r(\theta_0) = 0$ where r: $R^K \rightarrow R^{K-M}$ is twice continuously differentiable and its partial derivative matrix $R(\theta) = \partial r(\theta)/\partial\theta$ has rank $K - M$ for $\theta \in R^K$. Let the alternative hypothesis be $r(\theta_0) \neq 0$. Then let $\hat{\theta}$ and $\bar{\theta}$ be GMM estimates under the null hypothesis (H_0) and the alternative hypothesis (H_1), respectively, and let $V_{J+I}(\hat{\theta})$ and $V_{J+I}(\bar{\theta})$ converge to V in probability (with $V^{-1} = E[m_{ji}(\theta_0)m_{ji}(\theta_0)']$) under the null and the alternative hypotheses, respectively. Then the minimum distance statistic *D* and the Wald statistic *W* are defined as follows.

(4.1) $$D = (J+I)\{G_{J+I}(\hat{\theta})V_{J+I}(\hat{\theta})G_{J+I}(\hat{\theta}) - G_{J+I}(\bar{\theta})V_{J+I}(\bar{\theta})G_{J+I}(\bar{\theta})\},$$

and

(4.2) $$W = (J+I)r(\bar{\theta})'[R(\bar{\theta})\Lambda_{J+I}(\bar{\theta})R(\bar{\theta})']^{-1}r(\bar{\theta}),$$

where $\Lambda_{J+I}(\bar{\theta})$ is an estimator of the asymptotic variance of $\bar{\theta}$. Then both *D* and *W* are asymptotically distributed by a χ^2 distribution with degrees of freedom equal to $K - M$ (the number of additional independent restrictions under the null hypothesis). The results of the following hypothesis testing are summarized in table 3.

Based on these test statistics, we test the spillover effects in learning by doing. The spillover effects of learning by doing have been considered a possible source of sustained growth

---

markets, and thus the bell-shaped demand pattern over the product cycle (implicitly) reflects this interin the literature of the endogenous growth theory (see Lucas (1988); Stokey (1986, 1988); and Young (1991, 1993)). We first test the intergenerational spillover effects. With possible intergenerational spillover effects, we now measure the production experience of firm j, say E_{jt}, as: $E_{jt} = B_{jt} + \alpha_1 F_{jt}$, where F_{jt} is the cumulative outputs of the firm's previous generational DRAM, and α_1 is the intergenerational spillover rate ($0 \leq \alpha_1$). In the test, the null hypothesis is that $\alpha_1 = 0$. Hence, the degree of freedom of this hypothesis test is equal to 1. Both test statistics accept the null hypothesis (and thus reject the significance of intergenerational spillover effects) at the significance level of 0.05 in the 1M DRAM case. In the 256K DRAM case, the Wald statistic accepts the null hypothesis at the significance level of 0.05, but the minimum distance statistic rejects the null hypothesis at the significance level of 0.05. In other words, at the significant level of 0.05, the minimum distance statistic supports significant intergenerational spillover effects in the 256K DRAM. However, even in this case, the intergenerational spillover rate α_1 is not precisely estimated (see table 4). Next, we test whether learning effects spill over across firms. With possible industry-wide spillover effects, we measure the production experience of firm j as: $E_{jt} + \alpha_2(E_{gt} - E_{jt})$, where E_{gt} is the industry-wide production experiences, and α_2 is the industry-wide spillover rate ($0 \leq \alpha_2 \leq 1$). In this test, the null hypothesis is that $\alpha_2 = 0$, and the degree of freedom is equal to 1. As indicated in table 3, the spillover effects across firms are not supported in the learning process of the 256K DRAM or the 1M DRAM manufacturing.[22]

These hypothesis testing results for the spillover effects seem to be appealing in the DRAM case, considering that learning-by-doing effects are generated by engineers' adaptation to specific process and equipment through repeated operations while the actual physical design of the DRAM chip's internal structure and the fabrication facility vary across firms.[23] In addition,

---

generation substitution effects.

[22] It is interesting to note that Irwin and Klenow (1994) found no significant intergenerational spillover effects for the entire product cycle but significant industry-wide spillover effects.

[23] Hence, learning effects may not be carried over by the engineers with experience in other DRAM firms. The other possibility, however, is that the labor mobility in the DRAM industry is not substantial or at least

although a firm may try to utilize the previous generational production experiences by building a similar fabrication facility, there may still exist specific and idiosyncratic aspects of a new generational process and equipment. Our mixed test results on the intergenerational spillover effects for the 256K DRAM indicate that the previous generational production experiences may be effectively utilized in the mature stage when engineers adapted to specific and idiosyncratic aspects of a new generational process and equipment. As indicated in Irwin (1996), however, successful entries of new firms in this industry may not be supportive of the significance of intergenerational spillover effects. On the other hand, we can consider the possibility that the firms which produce with little experience of the previous generation may tend to have a good unobservable (productivity shock) while the firms with lots of experience of the previous generation may tend to produce despite a poor unobservable. However, in our data, we do not have serious this kind of selection bias. In the 1M DRAM case, all the firms in our data had some production experience of the previous generation while only 3 firms (out of 19 firms) had no production experience of the previous generation in the 256K DRAM case. Even eliminating these 3 firms from our data, we reached the same conclusion.

We now proceed to test whether an initial inefficiency level of the production process (c_j in the marginal cost function specified in (3.4)) may differ across firms. An initial inefficiency level of the production process of a firm may depend on the advancement of the firm's fabrication equipment. To see if later-entering firms had any advantage of employing a better fabrication equipment, we specify that $c_j = c + c_1 a_j$, where a_j is the difference between the quarter when firm j began to produce and the quarter when the first DRAM of the generation was introduced to the market. Then, based on the test statistics defined in (4.1) and (4.2), we accept the null hypothesis that $c_1 = 0$. We have also used alternative specifications, $c_j = c + c_1\ln(a_j+1)$ and $c_j = c + c_1\exp(a_j)$, and reached the same conclusion.

---

was unusually low during the STA period. But we are not aware of any evidence that the labor mobility was unusually low during the STA period.

Recall that the firm-specific P-MC ratio is assumed to be: $\lambda_{jt} = \lambda_t(\lambda_0 + f(B_{jt}))$ in equation (3.2) of section 3. As a baseline case, we have employed: $f(B_{jt}) = \ln(B_{jt}+1)$,[24] in which the firm-specific P-MC ratio increases with the cumulative production at a *decreasing* rate. Hence, we also tried two alternative specifications such as $f(B_{jt}) = \exp(B_{jt})$ and $f(B_{jt}) = B_{jt}$. However, both of these alternative specifications lead to huge values of the J statistic for the over-identification restrictions and an unreasonable negative sign of λ_t in some quarters during the STA period.[25] Based on $\lambda_{jt} = \lambda_t(\lambda_0 + \ln(B_{jt} + 1))$, we further test whether λ_0 can be normalized to be 1.[26] Then under this null hypothesis, λ_t can be interpreted as the P-MC ratio *for entrants*. We can accept this null hypothesis as indicated in table 3.

4.2. Estimation Results and Interpretations

Based on the above specification tests, we report our main estimation results in table 4. We also report the estimation results of the 256K DRAM case with intergenerational spillover effects. In our main estimation, we estimate 3 cost-side parameters and 20 dummy variables for $\{\lambda_t\}$ of the Japanese P-MC ratios with 31 instrumental variables (11 demand and cost shifters and 20 dummies variables for λ_t's), and thus the degrees of freedom are 8. J statistic turns out to be low enough to accept the over-identification restrictions at the significance level of 0.05 in the 1M DRAM case or at the significance level of 0.01 in the 256K DRAM case.

Table 4 indicates that all the cost parameters are significant except the logarithm of exchange rate (and the intergenerational spillover rate). The estimates of 'constant' indicate that

[24] We use $\ln(B_{jt}+1)$ instead of $\ln(B_{jt})$ since B_{jt} will be 0 for entrants. In addition, the coefficient of $\ln(B_{jt}+1)$ is normalized to be 1 since this coefficient cannot be identified from λ_t and λ_0.

[25] The negative sign of λ_t implies a negative P-MC ratio.

[26] To avoid that the P-MC ratio is equal to 0 for entrants, we set under the null hypothesis that $\lambda_0 = 1$ instead of 0.

the initial marginal production cost is a little bit lower in the 1 DRAM manufacturing than the 256K DRAM manufacturing. A key estimate of our model is the learning elasticity, γ. As well known, the slope of the learning curve, say *s*, is related to the learning elasticity by the formula: $s = 2^{\gamma} \times 100$, and (1 - *s*) indicates the percentage cost reduction when cumulative output is doubled (*learning rate*). The learning elasticity is estimated to be –0.43 for 1M DRAM and –0.48 for 256K DRAM, respectively. Hence, in the 1M (256K) DRAM manufacturing, the marginal cost falls by 25.8 percent (28.3 percent) if the cumulative output is doubled.[27] As indicated in Dick (1991) and Flamn (1993c), in the DRAM industry, it has been widely believed that the learning rates are between 25 and 30 percent.[28] The estimated learning rates reported in table 4 are higher than those estimated in Irwin and Klenow (1994) (18.4 percent and 19.6 percent, respectively). Note that in Irwin and Klenow (1994), production experiences are specified to include industry-wide spillover effects (but no intergenerational spillover effect). A more rapid learning rate, 36 percent, is reported on the manufacturing of the wide-bodied commercial aircraft in Benkard (2000a). However, note that Benkard (2000a) also found that about 60 percent of experiences are not carried over due to organization forgetting in the case of the wide-bodied commercial aircraft.

As discussed above, λ_t can be interpreted as the P-MC ratio of a Japanese entrant (if there is) in quarter *t*. As shown in table 4, all the estimates of λ_t are statistically significant but have relatively low values in 1986:4 – 1987:2, 1988:1 for the 256K DRAM and in 1986:4 – 1987:3 for the 1M DRAM. As discussed in section 3.1, until 1988, the DRAM prices of the domestic Japanese market were substantially lower than the worldwide average prices. Since the prices in the non-Japanese market were higher than the worldwide average prices during this period, the

[27] The intergenerational spillover rate of the 256K DRAM is (imprecisely) estimated to be 0.63. However, either with or without intergenerational spillover effects of the 256K DRAM, we do not find any significant difference in following quantitative analyses of Japanese and non-Japanese firms' mark-ups during the STA period.
[28] In addition, an anonymous source confirmed that the learning elasticity calculated from the internal data analogous to the cost data is close to the estimates reported in table 4.

estimates of λ_t (and thus the mark-ups of Japanese firms) during this period may have *downward biases* in our worldwide market data.

To analyze the extent to which the price floors and quotas set during the STA period affected the Japanese firms' behavior, we proceed to calculate each firm's mark-up based on the estimates of table 4. Figures 7 and 8 illustrate the (weighted and simple) average mark-ups of Japanese and non-Japanese producers for the 256K DRAM and the 1M DRAM during the STA period, respectively. The weighted average mark-up is calculated by each firm's mark-up multiplied by the firm's market share within the group. Without a big difference in productivity shocks, we expect bigger firms to have larger production experiences and lower marginal costs as well as smaller incentives to lower their mark-ups for learning-curve optimization. Hence, the weighted average mark-up is expected to be usually higher than the simple average mark-up. As expected, the weighted average was always higher in both generations of DRAMs. Overall, however, the gap between the weighted and the simple average mark-ups became smaller over the product cycle. In other words, aggressive, smaller, later-entering firms quickly caught up with the bigger, earlier-entering firms.

A noticeable characteristic of figures 7 and 8 is several sharp drops of the average mark-ups. These drops, as we expect, are caused by the *entries*. The entry usually caused a huge gap between the weighted and the simple average mark-ups as clearly shown in figures 7 and 8. This huge cap reflects that entering firms, despite their small market shares, usually set very aggressive, negative mark-ups to explore learning-curve optimization. In the case of the 256K DRAM, which was already in the mature stage with 16 incumbents in the beginning of the STA, we observe entries in 1987:1, 1988:1 and 1988:3. The unusual sharp drop of the Japanese average mark-up in 1987:1 reflects the entry of Sharp (a Japanese firm) accompanied by the unusual exit of AT&T. The deep drops of both Japanese and non-Japanese average mark-ups in 1988:1 coincided with the entry of one Japanese firm and two non-Japanese firms accompanied by the exit of one non-Japanese firm, ending up with the average of Japanese mark-ups lower than that of non-Japanese mark-ups. The huge drop of the non-Japanese simple average mark-up in 1988:3 indicates the aggressive entry of Goldstar (a Korean firm). In the case of the 1M DRAM, which was in the expanding stage of the product cycle, we observe more active entries in 1986:4, 1987:4, 1988:1, 1988:3, 1989:1, 1989:3, 1989:4 and 1990:1, which raised the number of (incumbent) firms from 8 to 19 during the STA period. The entries except in 1988:3, as illustrated in figure 8, caused sharp drops of the simple average mark-ups. Japanese firms entered in 1986:4, 1988:1 and 1989:1 while non-Japanese firms entered in 1987:4, 1988:1, 1988:3, 1989:4 and 1990:1. Figure 8 also indicates that the exit of AT&T in 1987:1 was accompanied by a sharp drop of the non-Japanese average mark-up in 1987:1, which implies that AT&T might have set a higher mark-up in the previous quarter. The drops of the Japanese simple average mark-up in 1987:2 and the non-Japanese simple average mark-up in 1990:4 indicate relatively big negative productivity shocks of smaller Japanese and non-Japanese firms, respectively. The extremely huge drop of the non-Japanese average mark-up in 1987:4 coincided with the entry of three non-Japanese firms. This extremely huge drop is mainly due to a rare coincidence of huge negative productivity shocks of two of these three firms in that entering quarter. The output levels of these two firms (Samsung and Siemens) were extremely low in this quarter compared to the quarters that follow.

The patterns of the average mark-ups of *unrestrained* non-Japanese firms in figures 7 and 8 also support the significance of learning-curve optimization in the DRAM industry. The (both weighted and simple) average mark-ups of unrestrained non-Japanese firms were much lower in the 1M DRAM case than in the 256K DRAM case. As calculated in table 5, the mean values of these weighted (simple) average mark-ups were –4.7 percent (–60.1 percent) for the 256K DRAM and –66.5 percent (–593.4 percent) for the 1M DRAM, respectively. Even after excluding the extremely low mark-ups in 1987:4, these mean values of the weighted (simple) average mark-ups numbers are –32.4 percent (–73.9 percent) for the 1M DRAM. These results are consistent with the learning-curve optimization behavior, considering that the 1M DRAM was in the earlier, expanding stage of the product cycle while the 256K DRAM was already in the mature stage during the STA period. Moreover, the trend of the average mark-ups of unrestrained non-Japanese firms in both generations of DRAM demonstrates counter-product-cyclic mark-ups consistent with the learning-curve-optimization behavior. In the 256K DRAM case, the weighted average mark-up became positive from 1990:4 while in the 1M DRAM case, the weighted average mark-up remained negative during the STA period.[29]

In contrast to the case of unrestrained non-Japanese firms, as indicated in table 5, the weighted average mark-ups of Japanese firms were *higher* for the 1M DRAM (13 percent) than the 256K DRAM (11.6 percent). This is a striking result, considering the different stages of the product cycle of these two generations of DRAM. Naturally, the difference between the Japanese and the non-Japanese (weighted and simple) average mark-ups was much bigger in the 1M DRAM case. To quantify the extent to which the price floors and quotas set during the STA period affected the Japanese firms' behavior in the 256K market and the 1M DRAM market, however, we need to further calculate, based on the estimates of table 4, both the myopic-optimization mark-ups and the learning-curve-optimization mark-ups (see (2.2) and (2.3)).

Figures 9 – 10 compare these simulated average mark-ups with the actual Japanese average mark-ups during the STA period. As discussed above, due to the nature of our data, the estimates of the Japanese mark-ups in the earlier period of the STA (1986:4 – 1988:4) might have had downward biases. Indeed, the estimated Japanese firms' mark-ups during the STA period turned out to be lower than the myopic-optimization mark-ups in 1984:4 – 1987:2 and 1988:1 in the 256K DRAM case and in 1986:4 – 1987:3 in the 1M DRAM case.[30] However, except these

[29] Based on the cost data collected by MicroDesign Resources (MDR), Aizcorbe (2002) calculated Intel's margins in the microprocessor market. However, these cost data are collected when yield rates (or defect rates) are approaching or have reached maturity. Hence the margins calculated in Aizcorbe (2002) may not reflect the aggressive, low (or negative) margins to explore the learning effects in the earlier stage of product cycle.

[30] In these quarters (except 1987:4 in the 1M case), the Japanese average mark-ups were even lower than the learning-curve-optimization mark-ups. A possible explanation (if the downward biases are not sufficiently big) is that the Japanese DRAM producer were not restrained but set a negative indirect

quarters, the price floors and quotas set during the STA period were more than binding constraints of the learning-curve optimization: they facilitated *collusive behavior* of the Japanese firms even beyond the myopic optimization. To our knowledge, this is the first econometric evidence on the STA as a facilitating practice, which has been widely speculated as in Irwin (1988, 1996). As indicated in table 5, during the STA period, the Japanese firms' (weighted) mark-up was on average 127 percent higher than the myopic-optimization mark-up in the 256K DRAM case and 29 percent higher in the 1M DRAM case. However, compared with the learning-curve-optimization mark-ups, the Japanese firms' (weighted) mark-up was 6.1 times as high in the 256K DRAM case and 11.6 times as high in the1M DRAM case. Hence, we can infer that the Japanese firms were more collusive in the 1M DRAM market which was in the expanding stage of product cycle.

The simulated mark-ups reported in table 5, however, do not reflect changes of optimal output levels under different behavioral assumptions of Japanese firms (such as myopic optimization). We are currently working on solving for these optimal output levels under these different behavioral assumptions. Our intuition, however, is as follows. Under the myopic optimization assumption, Japanese firms' sales and the total sales are likely to increase (although non-Japanese output levels may decrease). These increased firm sales will reduce the firm's marginal cost more rapidly while these increased total sales will reduce the market price. Therefore, even after adjusting the optimal output levels, we may end up with the mark-ups similar to those reported in table 5.

The implications discussed above indicate that Japanese firms (especially 1M DRAM producers) had enjoyed very lucrative profit margins during the STA period. In another way, later-entering non-Japanese (especially 1M DRAM) firms benefited from the STA. During the STA period, non-Japanese firms were overall later-entering firms in both generations of DRAM.

dynamic part of the mark-up strategically in the earlier period of the STA. Then these mark-ups might have been closer to the optimal mark-ups of dynamic oligopoly in (2.2).

Table 5 as well as figures 7 and 8 indicates that these later-entering non-Japanese firms set aggressive mark-ups to fully explore learning-curve optimization, rapidly expanding their shares especially in the market for the 1M DRAM. This finding is consistent with the changes of the concentration ratios of these two generations of DRAM as discussed in section 3.1 (see figure 6). Hence, the STA facilitated the growth of new comers (although they were not the U.S. firms) in the DRAM industry and thus at least resolved the worry of the U.S. dependency on Japan for the DRAM supply.[31]

4.3. Alternative Specifications for the STA

In the above analyses, we assumed that the Japanese firm set a price-cost margin by adopting a simple formula of the P-MC ratio during the STA period. In this subsection, we will discuss the estimation results based on two alternative specifications assuming that the Japanese firm might be able to commit to a more sophisticated price-cost margin closer to its myopic or learning-curve optimization. As discussed in section 3.2, we assume that in each quarter, the regulator (the MITI) set the corresponding quotas to each Japanese firm, which credibly implemented these price-cost margins. Let λ_t^S and λ_t^D the deviations of the price-cost margin from the myopic-optimization margin and the learning-curve-optimization margin, respectively. Hence, under these two alternative specifications, the Japanese firm's price-cost margin is now either

[31] These rapid expansions of later-entering non-Japanese firms, especially Korean producers, might have turned out to be important in the R&D competitions for the following generations of DRAM. For instance, Samsung, a Korean producer, eventually became the leader of the innovations of new generations of DRAM. This successful entry, expansion and R&D leader story may be explained by the effects of production experiences on innovations as in Chang and Park (2002) rather than by the intergenerational spillover effects of learning by doing.

(4.3) $$p_t - mc_{jt} = -\frac{\partial p_t}{\partial q_{jt}} q_{jt} + \lambda_t^S,$$

or

(4.4) $$p_t - mc_{jt} = -\frac{\partial p_t}{\partial q_{jt}} q_{jt} + \phi E[\frac{\partial mc_{jt+1}}{\partial B_{jt+1}} | \Omega_t] + \lambda_t^D.$$

As discussed in section 3, Japanese firms might have had incentives to explore the learning-curve optimization on one hand and incentives to collude to achieve higher profits on the other hand. If the STA facilitated the collusive behavior even beyond the myopic optimization as reported in the previous subsection, we expect λ_t^S and λ_t^D to have a positive sign. On the other hand, if the Japanese firms were eager to explore the learning curve optimization, we expect λ_t^D to have a negative sign.

Using the same specifications on the demand function and the marginal cost function in section 3 and the conditional moment restriction in assumption 3, we can combine equation (4.3) (or (4.4)) and the non-Japanese firm's optimal margin in (3.3) to estimate the cost-side parameters and λ_t^S (or λ_t^D). As reported in tables 6 and 7, λ_t^S is estimated to have positive values except 1986:4 – 1987:3 and 1988:1 – 1988:2 in both generations of DRAM and, in addition, 1988:3 and 1989:1 – 1989:3 in the 1M DRAM case. On the other hand, λ_t^D is estimated to have positive values except four quarters of the earlier period of the STA (1986:4 – 1988:4) in the 1M DRAM case. Recall that in the earlier period of the STA, the estimates of λ_t^S and λ_t^D might have downward biases due to the nature of our data.

However, there are several reasons that the estimation results based on these alternative specifications are less appealing. First, the estimate of learning elasticity seems too high in the case of the myopic-optimization-margin specification of the 1M DRAM in which the implied learning rate is 40 percent (γ= -0.72). As discussed in the previous subsection, the learning rates are believed to be between 25 and 30 percent in the DRAM industry. Second, the estimates of λ_t^S and λ_t^D for both generations of DRAM turned out to be overall imprecise. Especially, the estimates of the 1M DRAM case under the specification of the learning-curve optimization turn out to have huge standard errors in several quarters. Lastly, J statistics for the over-identification restrictions cannot be accepted even at the significance level of 0.01 in the case of the myopic-optimization-margin specification for the 256K DRAM and in case of the learning-curve-optimization-margin specification for both generations of DRAM.

5. CONCLUSION

In the paper, we have empirically evaluated the extent to which the price floors and quotas set during the STA period affected the Japanese producers' behavior and the evolution of the markets for the 256K and the 1M DRAMs. To estimate two incentives faced by the Japanese firms for learning-curve optimization and collusive behavior under this institutional environment, we proposed the behavioral assumption on the Japanese firms' price-cost margins set during the SAT period, highlighting two key aspects of the procedure of firm-specific price floors: firm-specific price floors were set by voluntarily reported (projected) static costs, and then firm-specific quotas were set to implement these price floors. Our estimation results based on this behavioral assumption indicated that the price floors and quotas facilitated the collusive behavior of the Japanese producers even beyond the myopic optimization except in the earlier period of the STA although there may exist downward biases of the estimated mark-ups in this period of the STA due to the nature of the data employed in the paper. Especially, the 1M DRAM Japanese producers had enjoyed very lucrative profit margins during the STA period. On the other hand, later-entering non-Japanese firms fully explored learning-curve optimization and restrained

Japanese productions, rapidly expanding their shares in the worldwide market, especially in the market for the 1M DRAM. Our estimation results also suggest significant learning-curve-optimization behavior and counter-product-cyclic mark-ups but do not support any spillover effects of learning by doing in this industry.

For a complete welfare analysis of the STA, however, we need further to fully calibrate the dynamic structural model including the incorporation of the equilibrium behavior of entry and exit. Since we are mainly interested in estimating mark-ups and marginal production costs in the paper, we consider only incumbent firms' price-cost margins, implicitly assuming that firms make entry/exit decisions before the production decision as in Ericson and Pakes (1995). For the welfare evaluations of the STA, however, we have to calculate the equilibrium prices and sales as well as entries and exits without any restriction. For this purpose, based on the estimated marginal production costs of this paper and other calibration assumptions, we can extend the structural model of (2.1) to conduct counter-factual simulation studies of the evolution of the DRAM industry. Huang and Park (in process) conducts this counter-factual simulation study of the evolution of the 1M DRAM industry, which was most significantly affected by the STA. Adapting the computational method in Pakes and McGuire (1994), Huang and Park (in process) calculates Markov perfect equilibrium(a) of dynamic oligopoly in the 1M DRAM market without any restrictions as well as under the alternative strategic trade policies such as anti-dumping tariffs. The computed paths are compared to the historical path, i.e., the actual evolution of the DRAM industry during the STA period, providing welfare evaluations of this strategic bilateral trade agreement.

References

AIZCORBE, A. (2002): "Why are Semiconductor Prices Falling So Fast? Industry Estimates and Implications for Productivity Measurement," mimeo., Federal Reserve Board.

ATHEY, S., and SCHMUTZLER, A. (2001): "Investment and Market Dominance," *RAND Journal of Economics*, 32, 1-26.

BALDWIN R. (1994): "The Impact of the 1986 US-Japan Semiconductor Agreement," *Japan and the World Economy*, 6, 129-152.

BALDWIN, R., and P. KRUGMAN (1988): "Market Access and International Competition: A Simulation Study of 16K Random Access Memories," in *Empirical Methods for International Trade*, ed. by R. Feenstra. Cambridge: MIT Press.

BENKARD, C. L. (2000a): "Learning and Forgetting: The Dynamics of Commercial Aircraft Production," *American Economic Review*, 90, 1134-1154.

_______ (2000b): "Dynamic Equilibrium in the Commercial Aircraft Market," mimeo, Stanford University.

BERRY, S. and A. PAKES (2001): "Estimation from the Optimality Conditions for Dynamic Controls," mimeo., Yale University.

BERRY, S., J. LEVINSOHN, and A. PAKES (1999): "Voluntary Export Restrictions on Automobiles: Evaluating a Trade Policy," *American Economic Review*, 89, 400-430.

CABRAL M. B., and M. H. RIORDAN (1994): "The Learning Curve, Market Dominance, and Predatory Pricing," *Econometrica*, 62, 1115-1140.

CHANG, S. and S. PARK (2002): "Production Experience and Persistent Leadership in R&D Competition," mimeo., SUNY at Stony Brook.

DASGUPTA, P., and J. STIGLITZ (1988): "Learning-by-Doing, Market Structure and Trade Policies," *Oxford Economic Papers*, 40, 246 - 26.

DICK, A. R. (1991): "Learning by Doing and Dumping in the Semiconductor Industry," *Journal* *of Law and Economics*, 34, 133-159.

_______ (1992): "An Efficiency Explanation for Why Firms Second Source," *Economic Inquiry*, 30, 332-354.

DOHLMAN, P. A. (1993): "The U.S.-Japan Semiconductor Trade Arrangement: Political Economy, Game Theory, and Welfare Analysis," Ph.D. dissertation, Duke University.

ERICSON, R., and A. PAKES (1995): "Markov Perfect Industry Dynamics: A Framework for Empirical Work," *Review of Economic Studies*, 62, 53-82.

FEENSTRA, R. C. (1995): "Estimating the Effects of Trade Policy," in *Handbook of International Economics, Vol. III*. G. ed. by M. Grossman and K. Rogoff. Amsterdam: North-Holland, 1553-1595.

FLAMM, K. (1996): *Mismanaged Trade?* Washington, D.C.: The Brookings Institution Press.

________ (1993a): "Measurement of DRAM Prices: Technology and Market Structure," in *Price Measurements and Their Uses*, ed. by M. F. Foss, M. E. Manser and A. H. Young. Chicago: Univ. of Chicago Press.

_______ (1993b): "Forward Pricing versus Fair Value: An Analytical Assessment of 'Dumping' in DRAM," in *Trade and Protectionism*, ed. by T. Ito and A. O. Krueger. Chicago: Univ. of Chicago Press.

_______ (1993c): "Semiconductor Dependency and Strategic Trade Policy," *Brookings Papers on Economic Activity: Microeconomics*, 249-333.

FUDENBURG, D., and J. TIROLE (1983): "Learning-by-doing and Market Performance," *Journal of Economics*, 14, 522 - 530.

GHEMAWAT, P., and M. SPENCE (1985): "Learning Curve Spillovers and Market Performance," *Quarterly Journal of Economics*, 100, 839 - 852.

GRUBER, H. (1992): "The Learning Curve in the Production of Semiconductor Memory Chips," *Applied Economics*, 24, 885-894.

HANSEN, L. (1982): "Large Sample Properties of Method of Moments Estimators," *Econometrica*, 50, 1029-1054.

HATCH, N., and S. REICHELSTEIN (1997): “Learning Effects in Semiconductor Fabrication,” mimeo, UC at Berkeley.

HUANG, C., and S. PARK (in process): “The 1986 Semiconductor Trade Arrangement and the Evolution of 1M DRAM Industry,” mimeo., SUNY at Stony Brook.

IRWIN, D. A. (1998): “Antidumping: The Semiconductor Industry,” Brookings Trade Forum, Washington, D.C.: Brookings Institution, 173-200.

_______ (1996): “Trade Politics and the Semiconductor Industry,” in *The Political Economy of American Trade Policy,* ed. by A. O. Krueger. Chicago: Univ. of Chicago Press.

IRWIN, D. A., and P. J. KLENOW (1994): “Learning-by-Doing Spillovers in the Semiconductor Industry,” *Journal of Political Economy*, 102, 1200 - 1227.

KRISHNA, K. (1989): “Trade Restrictions as Facilitating Practices,” *Journal of International Economics*, 26, 251-270.

_______ (1987): “Tariffs versus Quotas with Endogenous Quality,” *Journal of International Economics*, 23, 97-112.

LUCAS, R. E., Jr. (1988): “On the Mechanism of Economic Development,” *Journal of Monetary Economy*, 22, 3-42.

MOOKHERJEE, D., and D. RAY (1991): “Collusive Market Structure Under Learning-By-Doing and Increasing Returns,” *Review of Economic Studies*, 58, 993-1009.

NEWEY, W., and D. MCFADDEN (1994): “Large Sample Estimation and Hypothesis Testing,” in *Handbook of Econometrics IV*, ed. by R Engle and D. McFadden. Holland: Elsevier.

PAKES, A., and P. MCGUIRE (1994): “Computation of Markov Perfect Nash Equilibria I: Numerical Implications of a Dynamic Product Model,” *RAND Journal of Economics*, 25, 555-589.

PARK, S. (forthcoming): “Semiparametric Instrumental Variables Estimation,” *Journal of* Econometrics.

SPENCE, A. (1981): “The Learning Curve and Competition,” *Bell Journal of Economics*, 12, 49 - 70.

STOKEY, N. L. (1986): “The Dynamics of Industrywide Learning,” In *Essays in Honor of Kenneth J. Arrow, vol. 2, Equilibrium Analysis*, ed. by Walter P. Heller, Ross M. Starr, and David A. Starrett. New York: Cambridge Univ. Press.

_______ (1988): “Learning by Doing and the Introduction of New Goods,” *Journal of Political Economy*, 96, 701-717.

STOKEY, N., R. LUCAS, and E. PRESCOTT (1989): *Recursive Methods in Economic Dynamics*, Cambridge: Harvard University Press.

YOUNG, A. (1991): “Learning by Doing and the Dynamic Effects of International Trade,” *Quarterly Journal of Economics*, 106, 369-405.

______ (1993): “Invention and Bounded Learning by Doing,” *Journal of Political Economy*, 101, 443-472.

Table 1: Non-Japanese Firms' Market Shares (%)

Year.Quarter	1M DRAM		256K DRAM	
	average	max.	average	max.
1986.4	5.7	10.6	2.4	7.5
1987.1	0.6	0.6	4.2	12.3
1987.2	1.1	1.1	5.0	14.6
1987.3	1.3	1.3	5.0	15.0
1987.4	0.4	1.6	5.4	16.4
1988.1	1.3	2.0	4.3	16.2
1988.2	2.6	4.2	4.4	15.6
1988.3	2.5	5.0	4.1	15.4
1988.4	3.3	7.4	4.2	14.8
1989.1	3.3	7.5	4.1	13.2
1989.2	4.1	10.6	4.2	12.6
1989.3	4.2	13.0	4.2	12.6
1989.4	4.4	13.6	4.4	12.3
1990.1	4.9	15.7	5.4	12.3
1990.2	4.7	14.5	5.1	12.9
1990.3	4.8	14.5	5.3	13.6
1990.4	5.4	15.8	5.4	14.2
1991.1	6.0	15.1	5.6	14.5
1991.2	6.2	13.2	6.2	16.5
1991.3	6.3	12.9	4.6	10.9

Table 2: Demand and Cost Shifters

Demand Shifters
Aggregate quarterly GDP of the OECD countries
Quarterly changes of this aggregate GDP
Vintage of a generation of DRAM
Vintage squared
Vintage cubed
Vintage to the power four
Quarterly dummy variables

Cost Shifters
Exchange rates

Table 3: Specification Tests

	Hypothesis testing		Test statistics*		Conclusion**
Test 1	H0: No intergenerational spillover effect	265K DRAM	D = 8.81	W = 0.25	Accept H0***
	H1: not H0	1M DRAM	D = 0****	W = 0.01	Accept H0
			d.o.f = 1*****		
Test 2	H0: No industry-wide spillover effect	265K DRAM	D = 0	W = 0.08	Accept H0
	H1: not H0	1M DRAM	D = 0	W = 0.004	Accept H0
			d.o.f = 1		
Test 3	H0: No initial efficiency difference (c1=0)	265K DRAM	D = 0.06	W = 0.20	Accept H0
	H1: not H0	1M DRAM	D = 1.12	W = 3.29	Accept H0
			d.o.f = 1		
Test 4	H0: Normalization (ramda0 = 1)	265K DRAM	D = 0	W = 0.02	Accept H0
	H1: not H0	1M DRAM	D = 0	W = 0.0003	Accept H0
			d.o.f = 1		

* "D" is the minimum distance statistic and "W" is Wald statistic.
** The tests are conducted at the significance level of 0.05
*** However, the minimum distance statistic (D) rejects H0 at the significance level of 0.05.
**** D is set to be zero if we cannot find a lower value of the J statistic under H1
***** "d.o.f." means the degrees of freedom of the kai-sqaured distribution.

Table 4: Regression Results

Variables	1M DRAM		256K DRAM: Intergenerational Spillover Effects		256K DRAM: No Intergenerational Spillover Effects	
	Estimates	St. Errors	Estimates	St. Errors	Estimates	St. Errors
Cost-side Variables						
learning elasticity	-0.43 *	0.03	-0.48 *	0.04	-0.42 *	0.04
intergenerational spillover rate					0.63	1.27
ln(exchange rate)	-0.11	0.28	-0.05	0.08	-0.28 **	0.16
constant	4.60 *	0.29	4.70 *	0.38	4.31 *	0.61
ramda_t						
1986:4	0.09 *	0.02	0.07 *	0.02	0.09 *	0.03
1987:1	0.09 *	0.02	0.04 **	0.02	0.11 *	0.03
1987:2	0.09 *	0.02	0.07 *	0.02	0.09 *	0.02
1987:3	0.09 *	0.02	0.09 *	0.02	0.10 *	0.02
1987:4	0.13 *	0.03	0.10 *	0.02	0.10 *	0.02
1988:1	0.10 *	0.03	0.05 ***	0.03	0.06 **	0.03
1988:2	0.12 *	0.04	0.10 *	0.03	0.09 *	0.03
1988:3	0.16 *	0.05	0.12 *	0.03	0.11 *	0.03
1988:4	0.22 *	0.05	0.13 *	0.03	0.12 *	0.03
1989:1	0.18 *	0.06	0.15 *	0.04	0.13 *	0.03
1989:2	0.17 *	0.07	0.16 *	0.04	0.14 *	0.03
1989:3	0.16 *	0.05	0.13 *	0.03	0.11 *	0.02
1989:4	0.15 *	0.04	0.11 *	0.02	0.09 *	0.02
1990:1	0.14 *	0.03	0.12 *	0.03	0.10 *	0.02
1990:2	0.13 *	0.03	0.11 *	0.02	0.10 *	0.02
1990:3	0.12 *	0.03	0.10 *	0.02	0.09 *	0.02
1990:4	0.11 *	0.02	0.09 *	0.02	0.08 *	0.02
1991:1	0.10 *	0.02	0.10 *	0.02	0.09 *	0.02
1991:2	0.10 *	0.02	0.11 *	0.02	0.09 *	0.02
1991:3	0.10 *	0.02	0.11 *	0.02	0.09 *	0.02
elasticity of demand	1.80		1.80		1.80	
discount factor	0.95		0.95		0.95	
number of firms		7 to 19		16 to 19		16 to 19
number of observations		300		363		363
J statistic		11.5		18.0		9.2
degrees of freedom		8		8		7

*: significant at the significance level of 0.05.
**: significant at the significance level of 0.1.
***: significant at the significance level of 0.11.

Table 5: Mean Values of the Average Mark-ups

		1M DRAM		256K DRAM	
		weighted average	simple average	weighted average	simple average
Actual mark-ups during the STA period	Japanese				
	STA period	13.0%	-23.9%	11.6%	-14.5%
	1986:4 - 1988:4	-15.8%	-82.4%	-13.7%	-71.0%
	1989:1 - 1991:3	36.6%	24.0%	32.3%	31.7%
	Non-Japanese				
	STA period	-66.5%	-593.4%	-4.7%	-60.1%
		[-32.4%]*	[-73.9%]*		
	1986:4 - 1988:4	-144.1%	-1264.0%	-12.1%	-129.1%
		[-72.7%]*	[-113.8%]*		
	1989:1 - 1991:3	-3.0%	-44.8%	1.4%	-3.6%
Simulated Japanese mark-ups	Myopic optimization				
	STA period	10.1%	4.9%	5.1%	3.5%
	1986:4 - 1988:4	15.5%	6.9%	5.8%	4.0%
	1989:1 - 1991:3	5.7%	3.3%	4.6%	3.1%
	Learning-curve optimization				
	STA period	1.1%	-85.4%	1.9%	-10.1%
	1986:4 - 1988:4	-1.5%	-161.3%	-0.4%	-24.2%
	1989:1 - 1991:3	3.3%	23.3%	3.7%	1.5%

* excluding the non-Japanese average mark-up of 1987:4 in the 1M DRAM

Table 6: Alternative Regressions for 256K DRAM

Variables	Myopic optimization		Learning-curve optimization	
	Estimates	St. Errors	Estimates	St. Errors
Cost-side Variables				
learning elasticity	-0.50 *	0.04	-0.51 *	0.04
ln(exchange rate)	-0.08	0.09	0.09	0.08
constant	4.94 *	0.38	4.90 *	0.39
deviations from focal levels				
1986:4	-0.40	0.31	0.59	0.36
1987:1	-0.57	16.07	0.43	14.95
1987:2	-0.63	0.42	0.56	3.94
1987:3	-0.14	0.18	0.14	0.68
1987:4	0.02	0.13	0.45 *	0.14
1988:1	-0.02	14.69	0.62	14.44
1988:2	-0.15	0.38	0.47	2.63
1988:3	0.08	0.26	0.43	0.32
1988:4	0.21	0.21	0.62 *	0.18
1989:1	0.31 **	0.18	0.27 **	0.15
1989:2	0.40 *	0.16	0.39 *	0.13
1989:3	0.13	0.14	0.21 **	0.12
1989:4	0.07	0.13	0.70 *	0.21
1990:1	0.17	0.12	0.18 **	0.11
1990:2	0.15	0.12	0.59 *	0.17
1990:3	0.06	0.12	0.45 *	0.15
1990:4	0.04	0.11	0.61 *	0.19
1991:1	0.06	0.11	0.48 *	0.15
1991:2	0.06	0.11	0.29 *	0.11
1991:3	0.08	0.11	0.21 *	0.10
elasticity of demand	1.80		1.80	
discount factor	0.95		0.95	
number of observations	363		363	
J statistic	20.7		93.6	
degrees of freedom	8		8	

*: significant at the significance level of 0.05.
**: significant at the significance level of 0.1.

Table 7: Alternative Regressions for 1M DRAM

Variables	Myopic optimization		Learning-curve optimization	
	Estimates	St. Errors	Estimates	St. Errors
Cost-side Variables				
learning elasticity	-0.72 *	0.06	-0.56 *	0.04
ln(exchange rate)	1.44 *	0.28	0.44	0.28
constant	5.34 *	0.43	5.30 *	0.33
deviations from focal levels				
1986:4	-1.47	77.47	-0.07	0.93
1987:1	-1.70	8.58	1.34 *	0.56
1987:2	-1.10	52.88	0.14	19.53
1987:3	-2.24	4.18	-0.25	100.00
1987:4	1.21	1.46	0.43	134.77
1988:1	-2.06	53.14	-1.86	2e^3 ***
1988:2	-0.29	9.07	0.36	28e^3
1988:3	-0.10	4.74	-0.48	29e^3
1988:4	3.71 *	0.45	0.31	11e^3
1989:1	-2.36	31.62	0.69	219e^3
1989:2	-0.91	5.89	1.17	3e^6
1989:3	-1.47	1.96	1.27	15e^6
1989:4	1.64 *	0.52	1.01	3e^6
1990:1	1.46 *	0.23	1.09	836e^3
1990:2	1.27 *	0.23	0.93	188e^3
1990:3	1.16 *	0.19	0.70	30e^3
1990:4	1.06 *	0.15	0.62	6e^3
1991:1	0.95 *	0.14	0.57	1e^3
1991:2	0.91 *	0.12	0.49	127.04
1991:3	0.86 *	0.10	0.52	10.43
elasticity of demand	1.80		1.80	
discount factor	0.95		0.95	
number of observations		300		300
J statistic		16.5		30.0
degrees of freedom		8		8

*: significant at the significance level of 0.05.
**: significant at the significance level of 0.1.
***: 2e^3 = 2*1000.

Figure 1: Sales

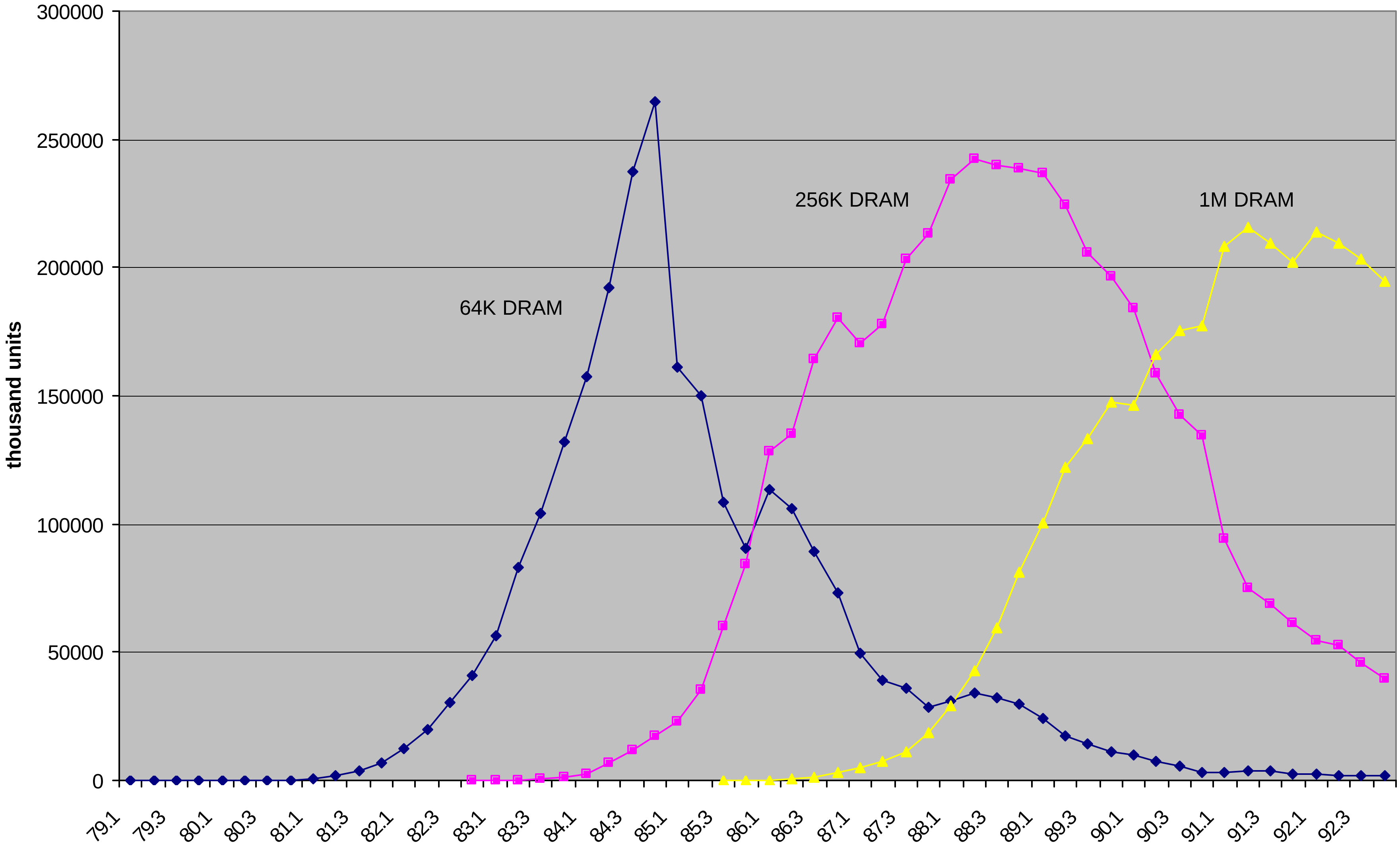

Figure 2: Entry

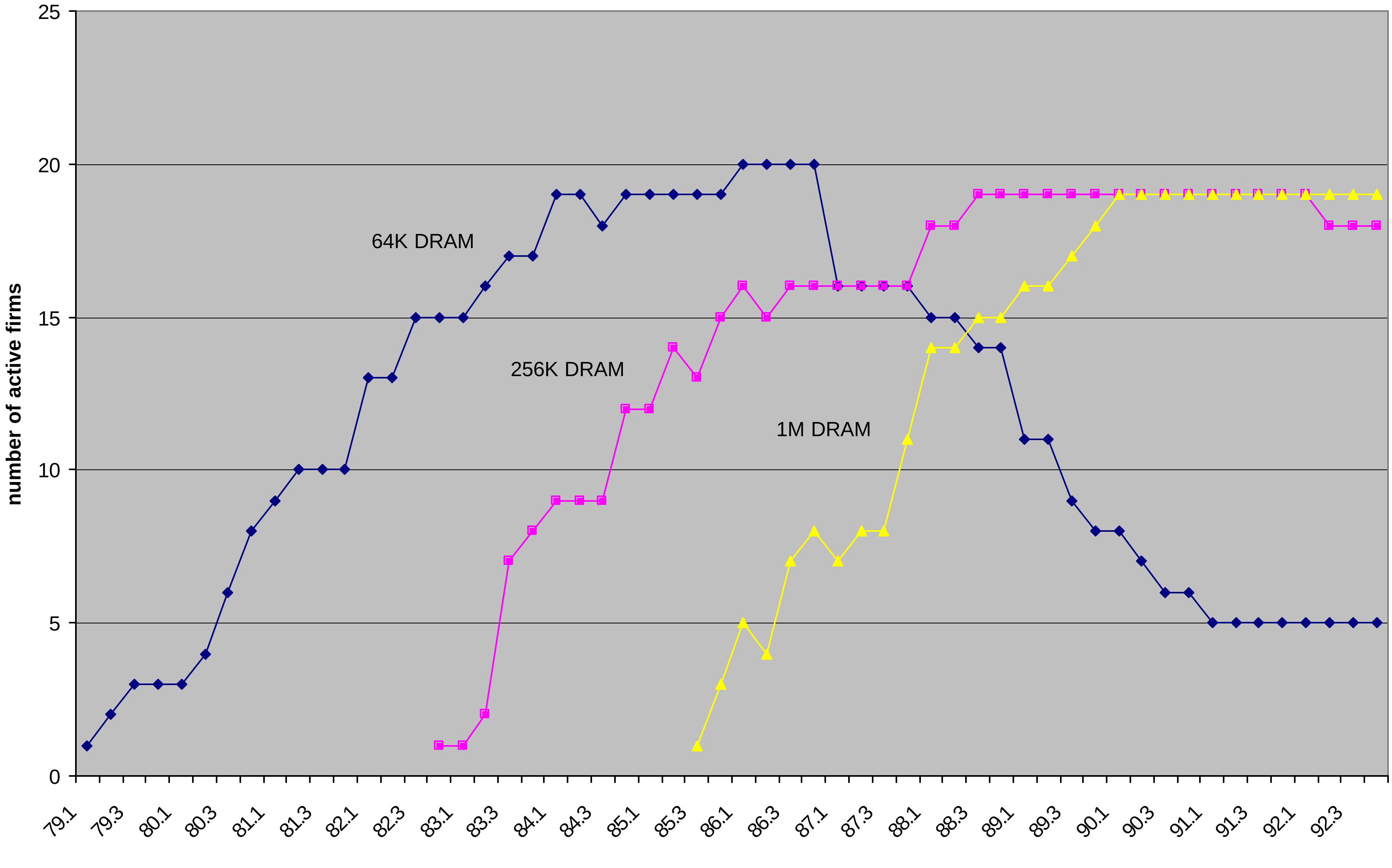

Figure 3: Herfindahl-Hirschman Index (HHI)

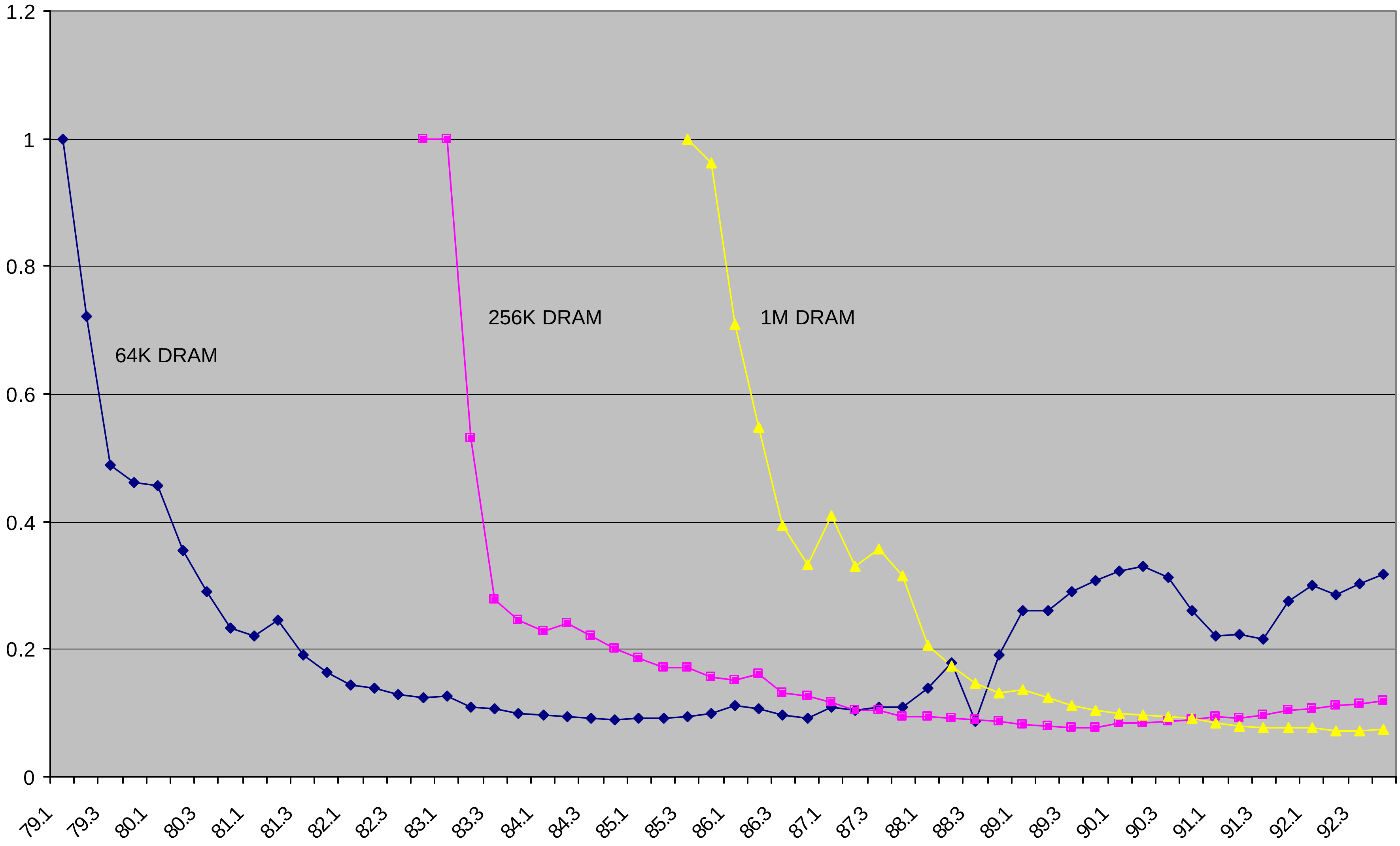

Figure 4: Prices of 4 genereations of DRAM

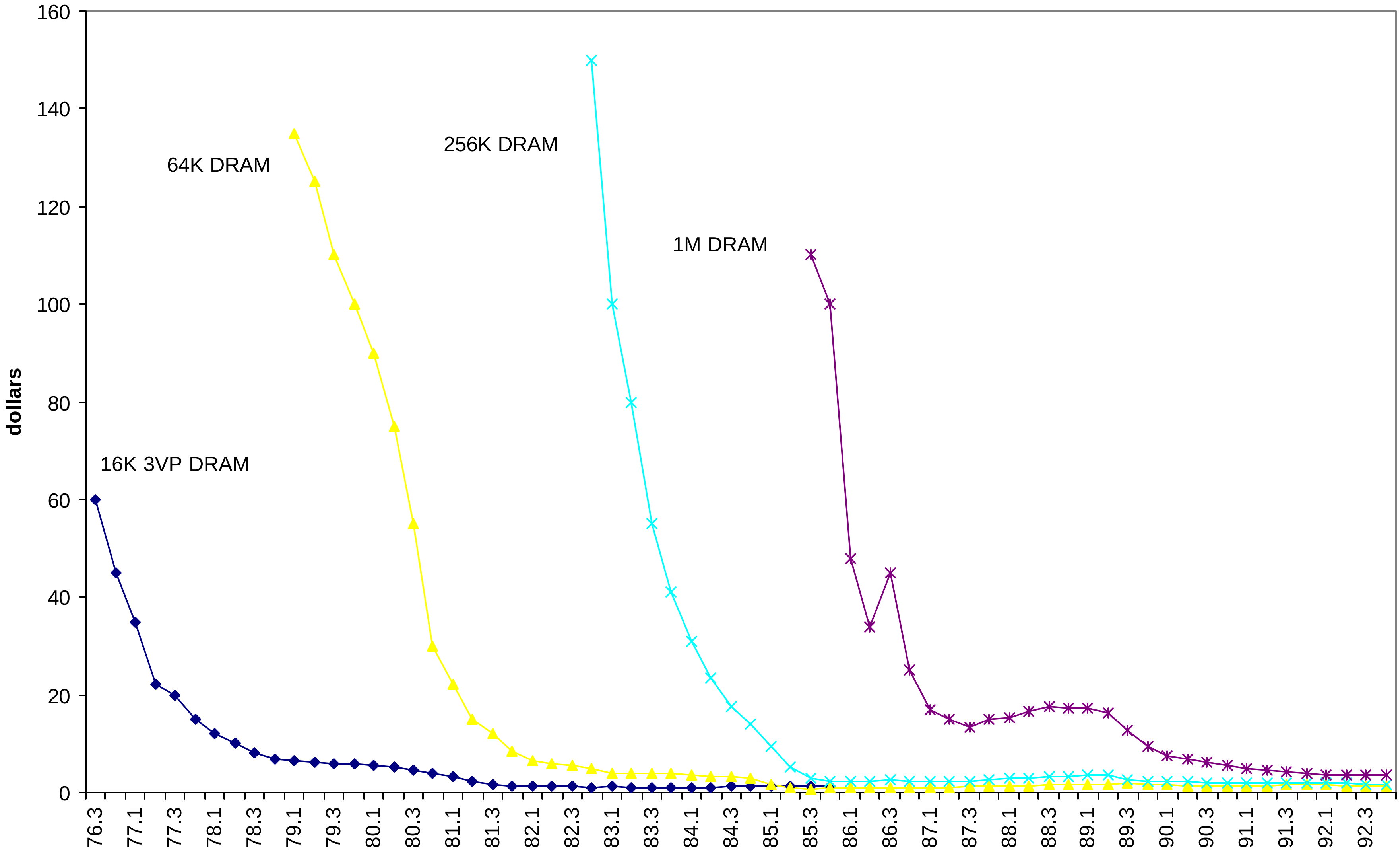

Figure 5: ln(prices)

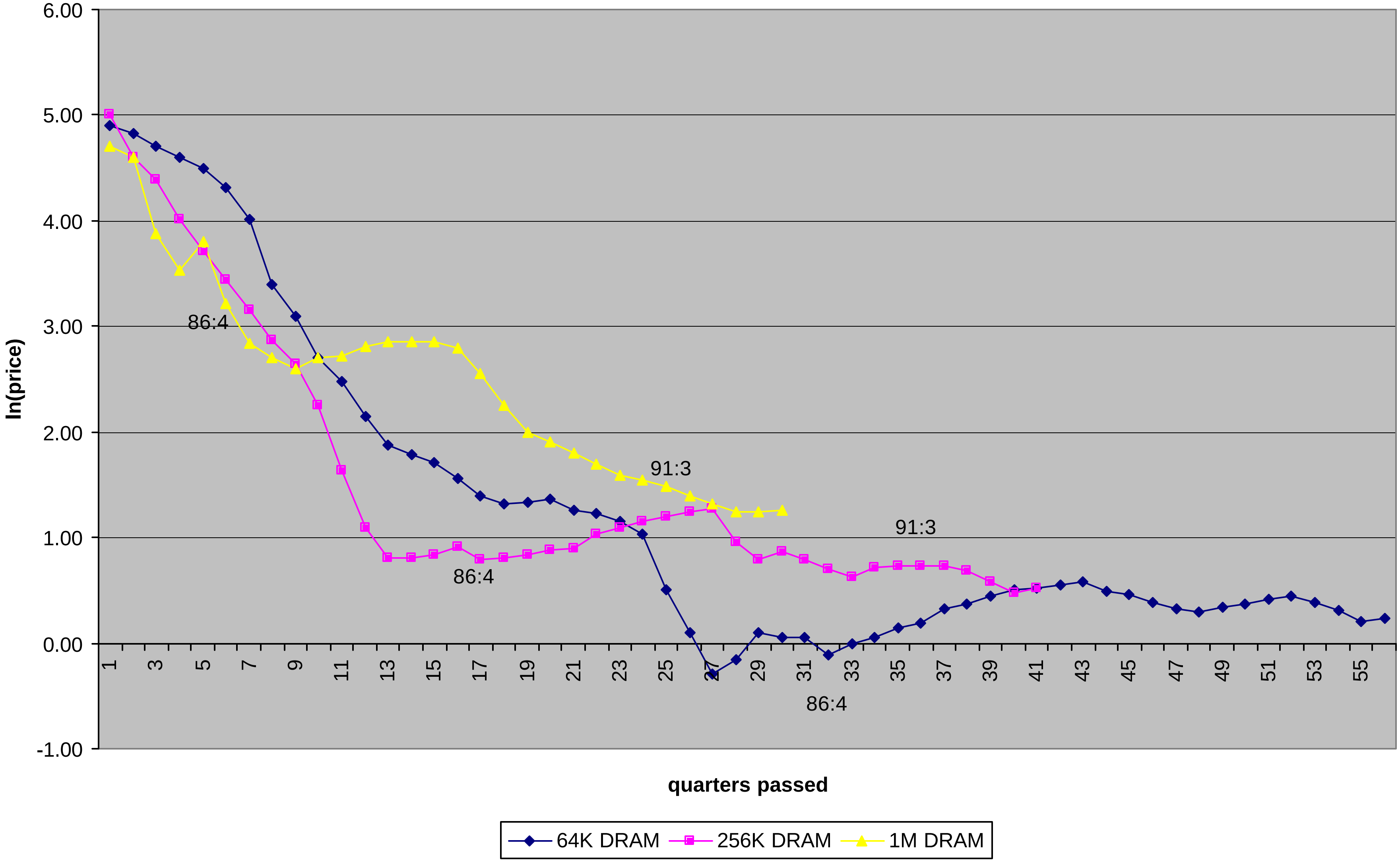

Figure 6: HHI in the earlier product cycle

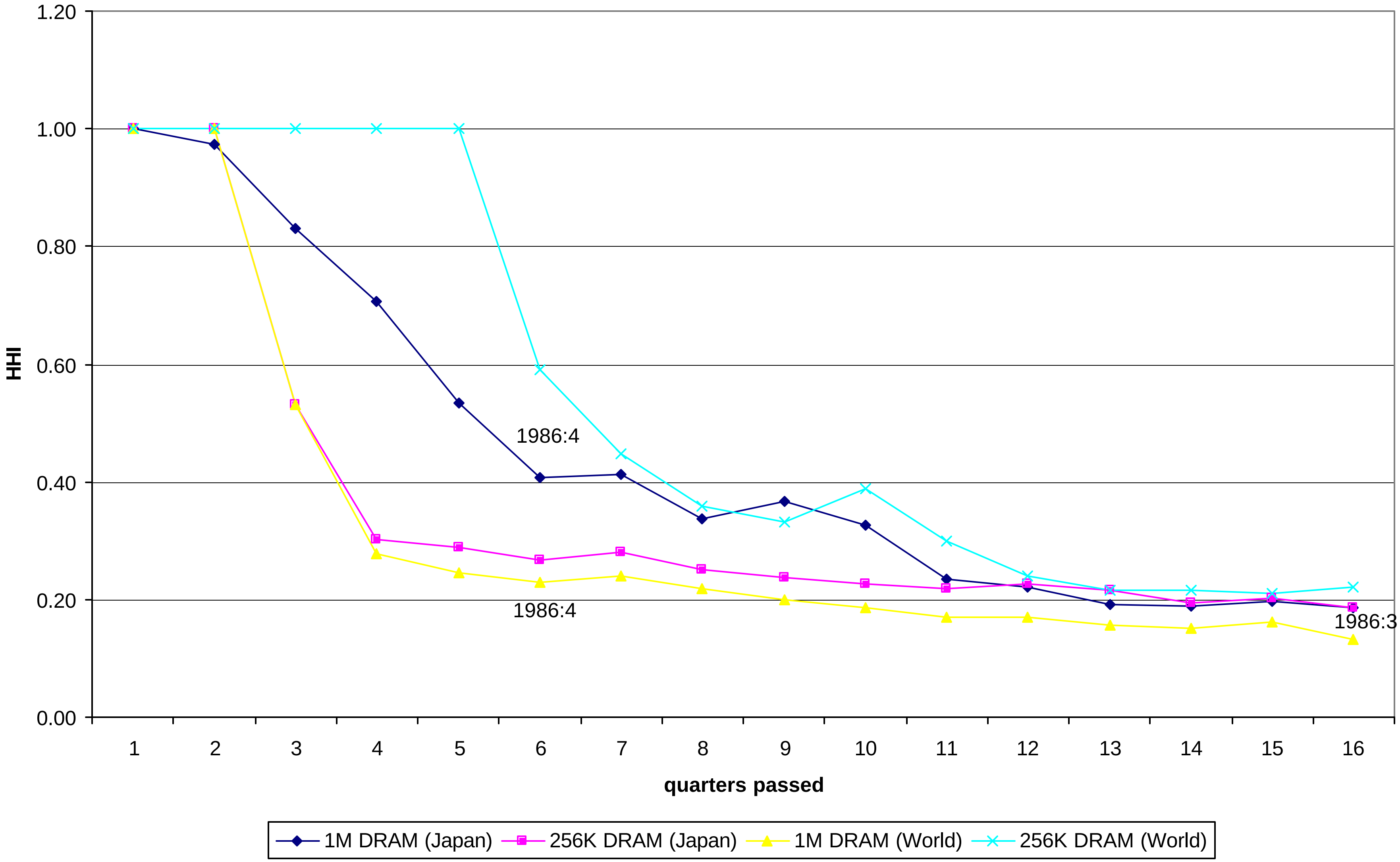

Figure 7: Mark-ups for 256K DRAM

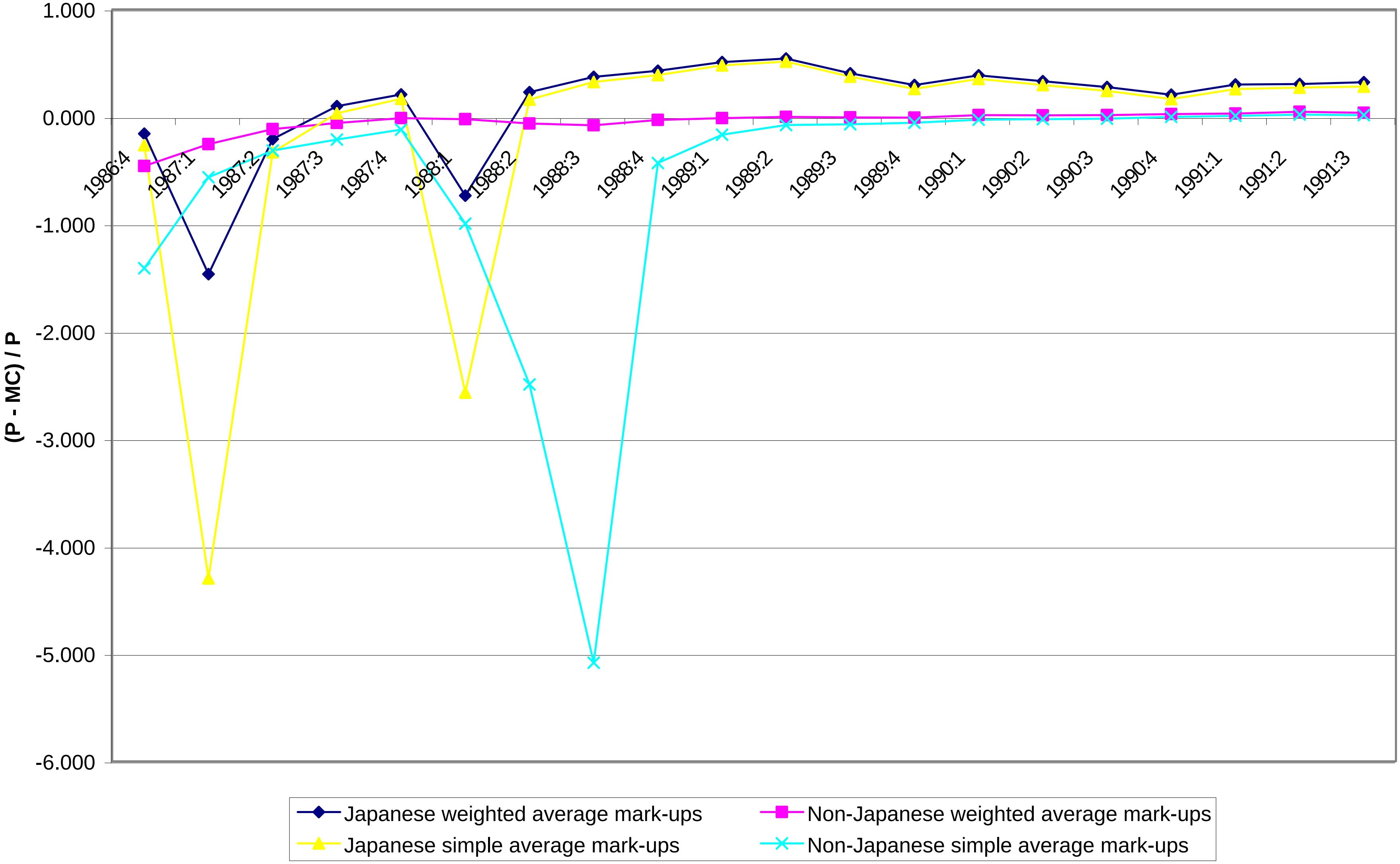

Figure 8: Mark-ups for 1M DRAM

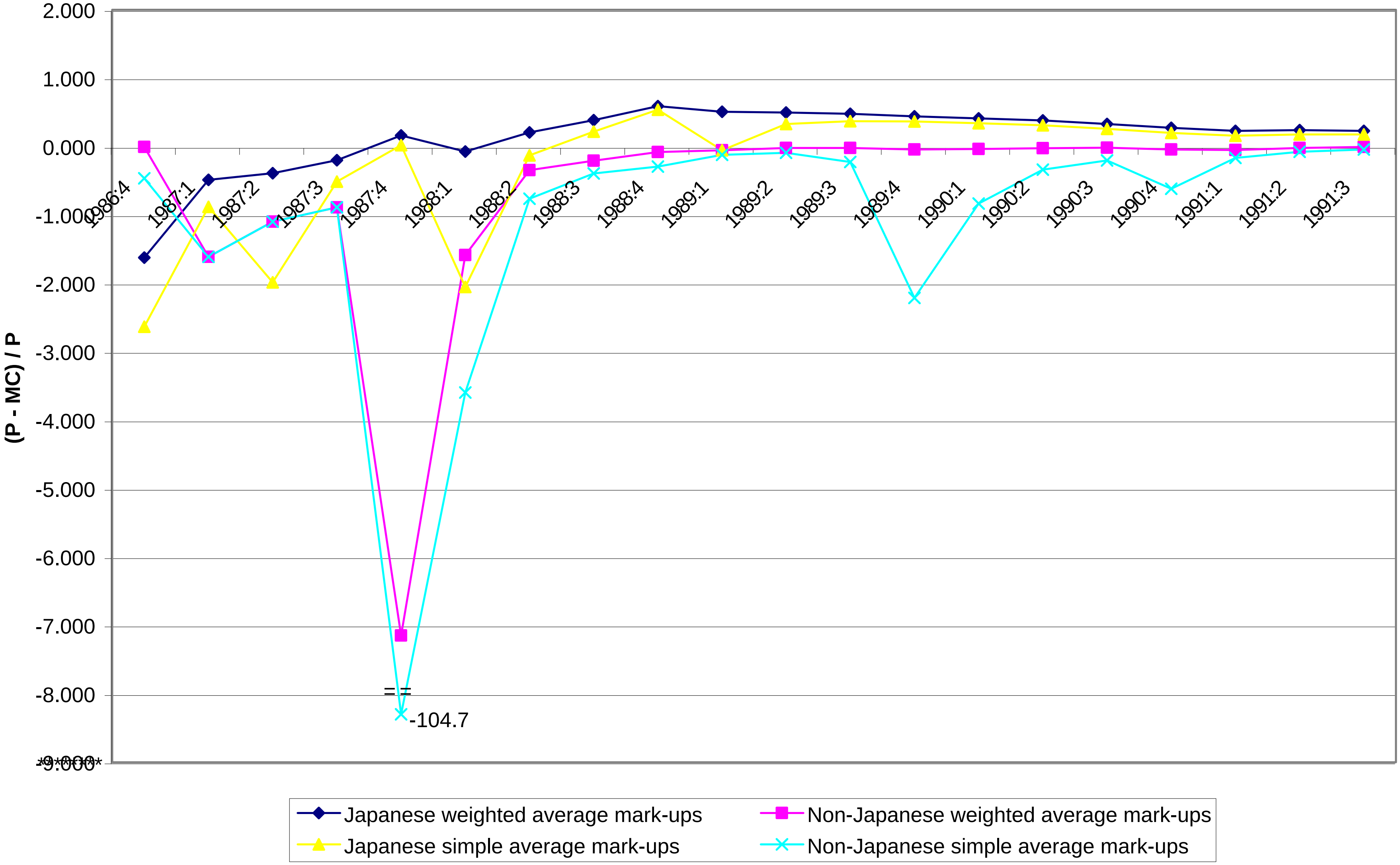

Figure 9: Japanese Weighted Average Mark-ups for 256K DRAM

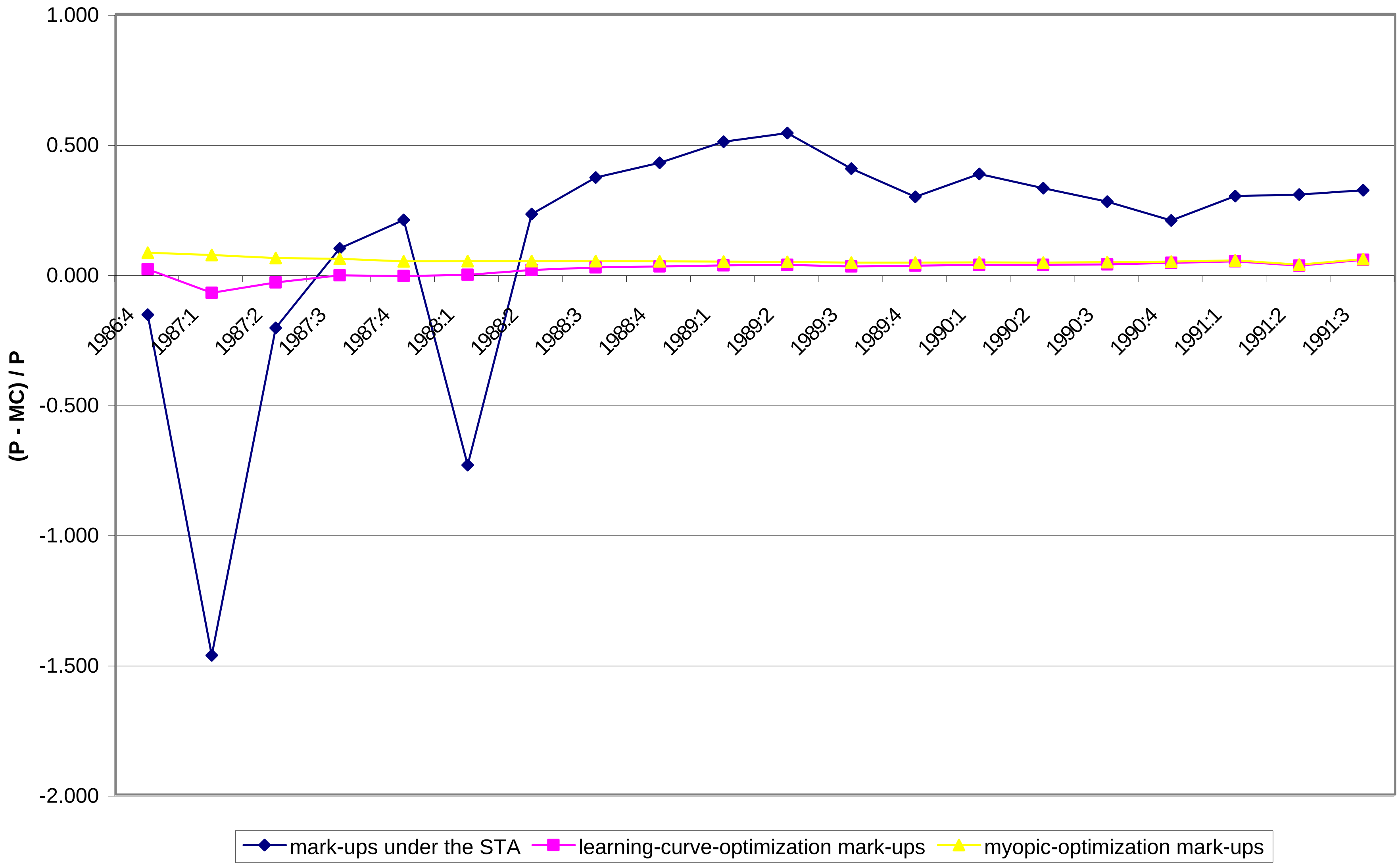

Figure 10: Japanese Weighted Average Mark-ups for 1M DRAMs