ECONOMIC ANALYSIS GROUP
DISCUSSION PAPER

Collusion Along the Learning Curve:
Theory and Evidence from the Semiconductor Industry

by

Danial Asmat*

EAG 16-4 August 2016

Revised July 2019

EAG Discussion Papers are the primary vehicle used to disseminate research from economists in the Economic Analysis Group (EAG) of the Antitrust Division. These papers are intended to inform interested individuals and institutions of EAG's research program and to stimulate comment and criticism on economic issues related to antitrust policy and regulation. The Antitrust Division encourages independent research by its economists. The views expressed herein are entirely those of the author and are not purported to reflect those of the United States Department of Justice.

Information on the EAG research program and discussion paper series may be obtained from Russell Pittman, Director of Economic Research, Economic Analysis Group, Antitrust Division, U.S. Department of Justice, LSB 9004, Washington, DC 20530, or by e-mail at russell.pittman@usdoj.gov. Comments on specific papers may be addressed directly to the authors at the same mailing address or at their e-mail address.

Recent EAG Discussion Paper and EAG Competition Advocacy Paper titles are available from the Social Science Research Network at www.ssrn.com. To obtain a complete list of titles or to request single copies of individual papers, please write to Regina Robinson at regina.robinson @usdoj.gov or call (202) 307-5794. In addition, recent papers are now available on the Department of Justice website at http://www.justice.gov/atr/public/eag/discussion-papers.html.

---

*Economic Analysis Group, Antitrust Division, U.S. Department of Justice, danial.asmat@usdoj.gov. The views expressed in this paper are those of the authors and are not purported to reflect the views of the U.S. Department of Justice. The empirical methodology for this article was constructed with the aid of publicly available media and academic sources while the author was a Ph.D. student at the University of Michigan, prior to any employment with the U.S. Department of Justice. The views expressed are not purported to reflect those of the U.S. Department of Justice.

Collusion Along the Learning Curve:

Theory and Evidence from the Semiconductor Industry

Danial Asmat*

July 31, 2019

Abstract

This paper studies the effectiveness of collusion in the DRAM cartel. Like other high technology products, DRAM is characterized by learning-by-doing and multiproduct competition. I hypothesize that collusion is more difficult to sustain on a new generation, where learning is high, than an old generation, where learning is low. A higher learning rate makes defection from a collusive equilibrium more attractive by reducing future cost. Empirical analysis exploits variation between cartelization and competition to estimate the change in firms' output decisions on each generation. Consistent with the hypothesis, cartel participants are estimated to cut output more on the oldest generation than newer generations. Output decisions on the newest generation also show evidence consistent with defection from collusive equilibria. Lastly, the paper presents a theoretical framework to analyze collusive equilibria with learning-by-doing and multiproduct competition. The model motivates various pieces of evidence that competition authorities can compile to guide antitrust investigations in high technology markets.

JEL Classification: D43, L13, L41, L63

*Economic Analysis Group, Antitrust Division, U.S. Department of Justice (danial.asmat@usdoj.gov). I am grateful to my advisors Maggie Levenstein, Valerie Suslow, Dan Ackerberg, Francine LaFontaine, and Brian Wu as well as Joe Harrington, Kai-Uwe Kühn, David Miller, Jagadeesh Sivadasan, Greg Werden, Nathan Wilson, and participants in several seminars, IIOC 2016, CRESSE 2016, and EARIE 2017 for comments. Data for this study was purchased with financial support from the Michigan Ross Doctoral Studies Office. I also thank Steve King of Gartner Inc., Jim Handy of Objective Analysis, and Sarim Shah of GlobalFoundries for numerous discussions about data and industry details. The empirical methodology for this paper was constructed with the aid of publicly available media and academic sources while the author was a Ph.D. student at the University of Michigan, prior to any employment with the U.S. Department of Justice. The views expressed are not purported to reflect those of the U.S. Department of Justice.

1 Introduction

Innovation is the foundational element of competition for a wide variety of high-technology products. It is particularly noteworthy that recent years have seen a steady discovery of price fixing cartels in dynamic, innovation-intensive industries: electrical auto parts, capacitors, LCD panels, cathode ray tubes, disk drives, and three different types of memory chips.[1] Some of these markets feature characteristics long known to enable collusion, including high barriers to entry and cooperative research and development. Yet they also share two additional characteristics less familiar in the academic literature on collusion. First, manufacturing displays learning-by-doing: a firm's cumulative output reduces the marginal cost of its future output. Second, firms steadily release new product generations based on technological advancements, such as those predicted by Moore's Law. Product generations with varying qualities therefore overlap on the market.

To study how firms collude in a high-technology industry with these two features, this article examines firm behavior during an illegal price fixing cartel in the Dynamic Random Access Memory (DRAM) market. Specifically, I hypothesize that collusion is more effective on older product generations than newer ones. To explore this hypothesis, consider that producing a unit of output on one product reduces its future cost. But the *rate* of cost decrease depends on whether the product is old or new. In an older product, because engineers have spent years engaging in trial-and-error, the learning rate is low. But in a newer product the learning rate is high. The faster is the learning rate, the more firms gain by defecting from collusion, and the more patient they must be to sustain it. These ideas are later formalized in a simple theoretical model.

Like numerous international cartels, the DRAM case also featured a fringe of smaller firms outside of the cartel agreement. The empirical model utilizes the presence of an exogenously competitive fringe to generate a testable hypothesis. The more effective is collusion on a product, the more the output of the fringe expands relative to the output of members. The key empirical hypothesis of the paper is that during the cartel, the fringe gains more share on the old generation than on newer generations.

DRAM production is a classic example of learning-by-doing: output allows firms to reduce cost-per-chip by reducing the rate of defective chips in manufacturing. The learning process repeats as chip generations are

[1] Fines from these cases are among the largest in U.S. antitrust history. AU Optronics, LG Display, and Samsung Electronics have each received penalties of $300 million or more since 2006: http://www.justice.gov/atr/public/criminal/sherman10.html.

released every two to three years, with long product life cycles creating overlapping generations. Moreover, the discovery and eventual antitrust litigation on the cartel has created evidence on its duration, participants, and non-participants.

To identify the output decisions of fringe versus cartel members, I construct a reduced form model that predicts firm output. The model exploits two levels of variation: conduct changed across time, and some firms joined the cartel while others did not. This permits a "difference-in-differences" estimate: the average change in members' output during the cartel period less the average change in non-members' output, separately for each generation.

Descriptive statistics show that the price of the 16Mb generation rose up to 25% during the cartel, whereas the price fell for the newer 64Mb and 128Mb generations. Consistent with the hypothesis, the DiD model estimates that cartel members reduced output by more on the 16Mb generation than the 64Mb and 128Mb generations, at a significance level of up to 1%. Although it is more challenging to estimate by how much cartel members reduced output on a single generation, the model's estimates can be used to construct a 95% confidence interval for total output reduction on the 16Mb generation between 3.1% and 24.6%.

To better understand the mechanism driving these results, I also formulate a straightforward test for defection from collusive equilibria in learning-by-doing industries. This test relies on the high correlation between each firm's order of entry into a generation and its subsequent market share ranking on that generation. If an unsuccessful attempt to collude stems from realized deviation by one or more firms, then there is basis for the correlation between order of entry and subsequent ranking to fall. This test shows that the correlation of 0.43 was significantly lower for the 128Mb DRAM generation than any other generation in the sample. The finding is consistent with the underlying hypothesis that collusion is most difficult to sustain on the newest generations during a cartel.

Literature Discussion This study's primary contribution is to generate empirical evidence from a cartel in a highly innovative market. It focuses on the classic problem of how cartels meet incentive compatibility constraints (Porter, 1983; Ellison, 1994; Levenstein, 1997; Scott Morton, 1997; Genesove and Mullin, 2001; Roller and Steen, 2006; Mariuzzo and Walsh, 2013; Clark and Houde, 2013, 2014, and Igami and Sugaya, 2017).[2] Many of these studies test canonical models of repeated games by exploiting the historical legality of cartels. Consequently, they are geared toward industries with straightforward supply functions, such as those in railroads, sugar, and cement.

In contrast, the present study examines a modern, high-technology product market in DRAM. It shows that a supply side dynamic can be the primary driver of incentive compatibility. de Roos (2004) uses a dynamic oligopoly model to explain the lysine cartel through the lens of entry deterrance and post-entry renegotiation. Similarly, this article shows that the dynamic incentives stemming from learning-by-doing should be considered in characterizing the DRAM cartel's behavior.

To assist in drawing more general insights, the paper also constructs a theoretical model of oligopoly centered on a two-period game. Following Siebert (2010) and Fudenberg and Tirole (1983), firms choose quantities in a learning period knowing that they gain a cost reduction as a function of output in a post-learning period.[3] This structure is then augmented to include vertically differentiated products and a collusive strategy. The model motivates various pieces of evidence that competition authorities can compile to guide horizontal investigations in high-technology markets.[4]

The remainder of the paper is outlined as follows. Section 2 and Section 3 describe the features of the industry, the cartel, and the data. Section 4 estimates the empirical model and tests for alternative hypotheses. Section 5 presents a model of collusion over the product life cycle, and summarizes its implications for antitrust policy. I conclude in Section 6 by discussing the limitations and future work.

2 Cartel Operations

Most cartel cases end before jury trial, limiting the amount of public information about their evolution. The DRAM cartel is noteworthy for leading to a federal criminal trial as well as civil class action trials. This section summarizes what the trial record implies about the cartel's timeline and its potential effectiveness, which are important to understand before proceeding to empirical analysis.

[2] See Levenstein and Suslow (2006) for a broader review of empirical literature examining incentive compatibility in cartels.

[3] Gardete (2016) also models firms as choosing quantities in the DRAM market. Other quantity-setting models of learning include Spence (1981), Ghemawat and Spence (1985), Dasgupta and Stiglitz (1988) and Cabral and Riordan (1997).

[4] Mookherjee and Ray (1991) also propose a model of collusion that includes scale economies and learning-by-doing, but only a single product. The present model emphasizes the interaction between learning and multi-product competition.

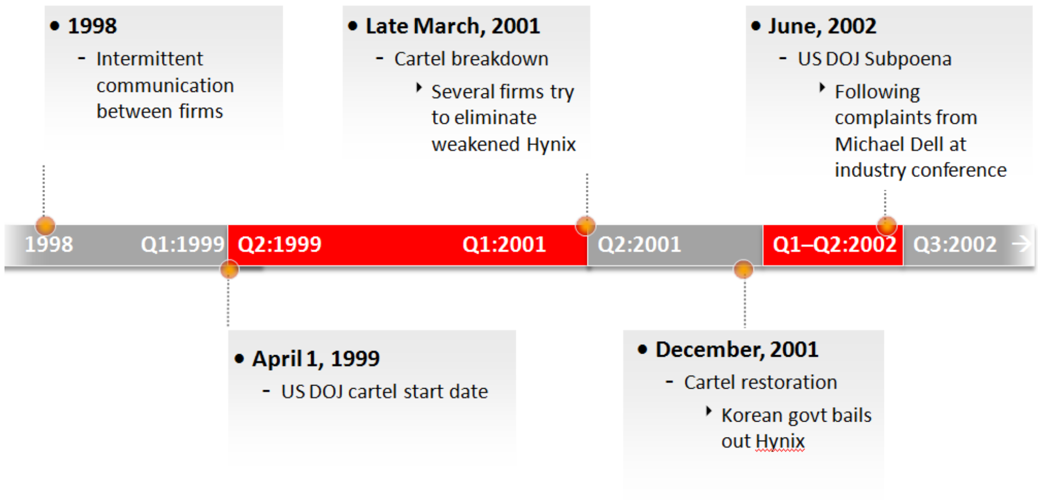


Figure 1: Events During Cartel Period
Sources: Noll (2014) and DOJ Press Releases

Figure 1 sequentially depicts the main events in the DRAM cartel. First, firm communication began in response to excess capacity. Next, cartel members cooperated to raise price from at least the second quarter of 1999 through the first quarter of 2001. From April through November 2001, however, they switched to active price competition in a bid to eliminate South Korean firm Hynix from the market. After state-owned banks bailed out the company, members reverted to the price fixing strategy until the DOJ issued a subpoena in June 2002 (Noll, 2014).[5]

These events illustrate three distinct phases of firm conduct: the cartel's first phase; the "Kill Hynix" phase; and the cartel's third phase. Empirical analysis will consider the dates corresponding to each phase as fixed. The "Kill Hynix" phase is treated as competitive, and the other two as collusive.

During its operation, cartel members communicated to raise price in three ways: specific price quotes to individual customers; general pricing levels across customers; and production volumes and inventories. Mid-level managers directed their sales representatives to communicate with rival counterparts in quoting prices to major OEMs. There is substantial email evidence documenting these price quotes, including a

[5] Communications between cartel members remained cooperative throughout these events, including the "Kill Hynix" period. For example, on September 20, 2001 executives from two other members discuss Hynix' position: 'On the assumption that Hynix gets new money, they would consider taking supply out of the market if others do the same.' Third Amended Class Action Complaint, MDL No. 1486, ¶ 214.

reference to a shared spreadsheet that tracked different firms' price targets. The evidence does not detail any specific monitoring or punishment strategy, or identify instances of cheating.

The scope to make strategic output decisions, including those factoring in future cost reductions, was likely limited to high-level executives. Senior executives held in-person meetings to discuss production figures and supply reduction. These meetings appear to have been organized between executives of two or more firms at a time rather than through a trade association or industry group. The evidence does not detail the content of the meetings, except where it was later memorialized through email, as in this example:

> *Mr. Chang confirmed that [firms] 'had reached an agreement ... to push up DRAM prices to US $3 a chip by stopping dumping.'* — Third Amended Class Action Complaint, ¶ 203-214

Official plea agreements state that the effectiveness of collusion varied by product:

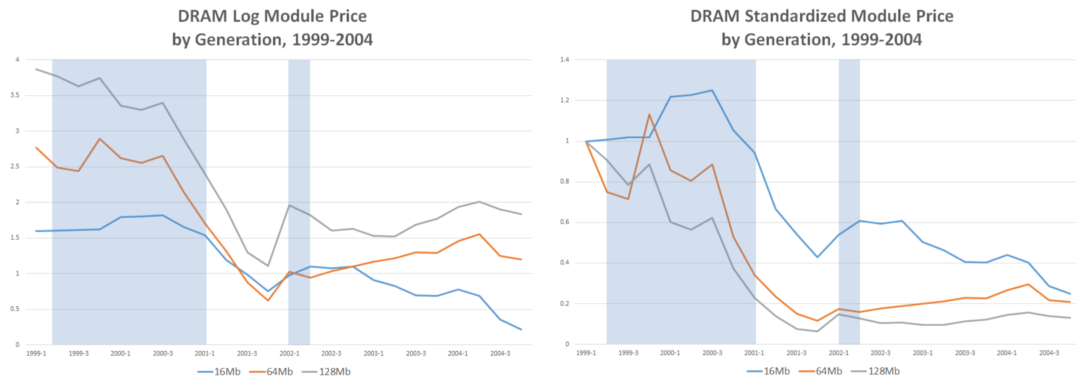
> *...at certain times during the Relevant Period, DRAM prices decreased significantly. Nevertheless, the Defendant, or its Corporate Founders, and their coconspirators reached agreements to limit the rate of price declines, which were achieved with varying levels of effectiveness.*
> — Plea agreement with [firm], March 22, 2006

The litigation record does not appear to link such statements to specific products or generations. The most suggestive remark comes from defendants' economic expert. Using price data from the market research firm IDC, the expert testified that prices during the third phase of the cartel rose significantly more for the 16Mb generation than newer ones. Testimony does not appear to analyze *why* such disparities may have arisen across generations, which motivates the present study.[6]

3 Data and Descriptive Evidence

I use a proprietary dataset from market research firm Gartner Research that lists quarterly shipments by firm and density generation, and quarterly market price by density generation, for all firms in the DRAM market from 1974-2011. Firm names allow accurate tracking of mergers, entry, and exit throughout the data sample. Separately, Gartner research reports list the quarterly worldwide PC shipments from 1997 onward. I consider the overlap of these two datasets, which spans from 1997-2011.

[6] The transcript of cross examination on this point is included in fig. D1.



(a) (b)

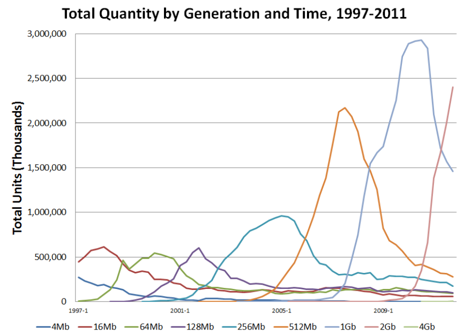
Figure 2: DRAM Price Trends: Cartel Active Periods Shaded

To gauge preliminary evidence, it is useful to consider the DRAM price during and after the cartel period. Figure 2 depicts this trend in two different ways. Panel (a) is the log of the module price, separately for each generation. Panel (b) standardizes the price level of each generation by its value in the the first quarter of 1999, prior to the official start of the cartel. During the first phase of the cartel, the price of the 16Mb chip rose up to 25% above its starting value. After the cartel ended, the 16Mb price declined, whereas the 64Mb and 128Mb price increased.

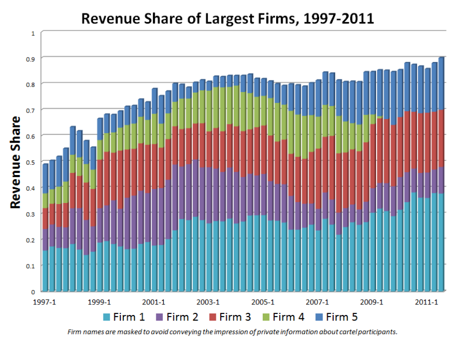
This pattern is consistent with a hypothesis in which the 16Mb price transitioned from a collusive to a competitive value after the cartel. However, price data suffers from small sample size because prices are available only at the generation-quarter level. There are 10 data points for cartel prices in each generation, and no firm-level variation. Consequently, I use firm-generation-quarter level shipments going forward.

Figure 3 represents several long-term trends in the DRAM market. Panel (a) depicts the total output (across all firms) by generation. Output takes an inverse U-shape for each generation and the total output of DRAM increased over time, consistent with increasing global demand for computers and electronics. Panels (b) and (c) highlight that the industry became steadily more concentrated over time. Panel (b) shows that the five largest firms accounted for an increasing share of industry revenues; (c) shows that with each new generation, some firms failed to advance (the three cartel generations are shaded in black). In the analysis that follows, I consider all active firms in each generation.[7]

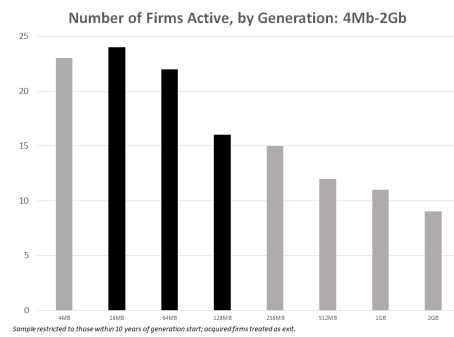
[7] Appendix C also displays the market share evolution of the largest firms separately for each of the three cartel generations.



(a) Total Output, by Generation



(b) Market Share, Select Firms



(c) Number Firms Active, by Generation

Figure 3

The empirical analysis focuses on firm output during the cartel period. Cartel members are identified by combining reports from the two largest antitrust authorities, the EC and DOJ, with additional firms named in the private class action complaint. This yields 11 cartel members, and the remaining firms are taken to comprise a competitive fringe. Table 1 summarizes the differences between cartel members and non-members from 1998 through the end of the cartel period, 2002.

Table 1: Firm Averages by Cartel Membership, 1998-2002

	N Firms*		Total Share†		Age (qtr)‡		Log Cum Output‡	
	1	0	1	0	1	0	1	0
16	11	11	0.76	0.24	26.2	20.2	12.5	10.6
64	10	9	0.90	0.10	14.0	8.9	11.4	8.9
128	10	6	0.97	0.03	7.7	4.8	9.7	6.7

* Sum across firms in group. † Mean across time, all firms in group
‡ Mean across time and firms
1 = cartel member; 0 = non-member.

Members have higher average shares, older ages, and more cumulative output than non-members. For this reason, the fringe is stronger the older is the generation. More generally, during the cartel period the 16Mb, 64Mb, and 128Mb generations accounted for 93% of total DRAM share by revenue.

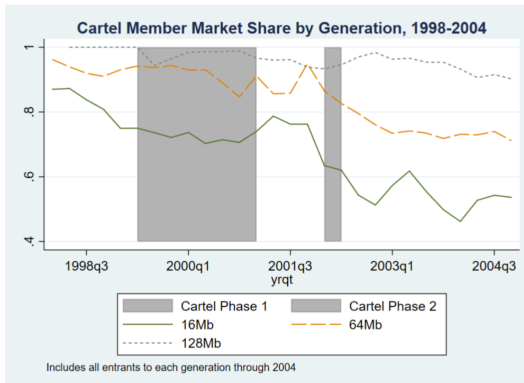
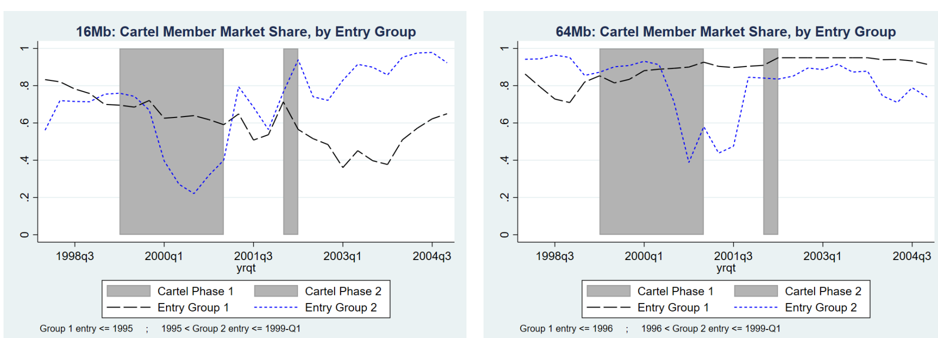


Figure 4: DRAM Cartel Member Share by Generation

Figure 4 plots the time series of total member share relative to all firms in the generation, before and after the cartel period. As reflected in the descriptive table, members have the largest share in the 128Mb generation, followed by 64Mb and 16Mb. In each generation, the cartel member share tends to decline over time, as smaller, later-arriving firms increase their production whereas larger firms eventually decrease it. This graph highlights the need to account for each firm's stage in the product life cycle of its generation.



(a) 16Mb

(b) 64Mb

Figure 5: Cartel Member Shares, by Generation and Entry Group

Share = $x_{et}(1-\alpha_{et})/\left[x_{et}(1-\alpha_{et})+y_{et}\alpha_{et}\right]$, where e=entry group, t=quarter

x=total member output, y=total non-member output, $\alpha = n_{members}\ /\ n_{all}$

To get a clearer picture, fig. 5 plots the cartel member share separately based on entry period for (respectively) the 16Mb and 64Mb generations. Each graph groups firms into two categories: early and later entrants. The cartel member share for each group is computed as the ratio of cartel member output to the sum of member and non-member output within the entry group.[8] Conditioning on entry period shows that, among the newer group of firms, member share reached its lowest point from 1998-2004 during the first phase of the cartel, and rebounded afterward. This effect is present in both generations. It is not driven by entry of new firms after the cartel started, as it excludes entrants to each generation on or after 1999-Q2.

The pattern in fig. 5 suggests that at least during the cartel's first phase, certain cartel members reduced output relative to outsiders on these two generations. However, the graph does not condition on cumulative output. It also ignores the role of firm decisions across product generations and of demand. To account for the effect of these variables, section 4 formulates a regression model that compares the output of firms in the cartel to those outside.

4 Empirical Results

4.1 Empirical Model

For each generation of interest, the research design estimates the cartel's effect on member output. This challenge is pursued by exploiting two sources of variation. Firm conduct is hypothesized to change across time: from competition to collusion, and vice versa. Firms are also hypothesized to differ in conduct: cartel participants and the competitive fringe.

Index firms by i, generations by k, and time by t. The panel data specification is given by a DiD model that accounts for the effect of accumulated output, cannibalization, and demand:

$$\begin{aligned} ln(q_{ikt}) = {} & \beta_0 + ln(x_{ikt})\beta_1 + \sum_{i=2}^{I}\Big[ln(x_{ikt}) \times \alpha_i\Big]\beta_i \\ & + ln\,(q_{irt})\,\omega_1 + ln(PC_t)\eta_1 + \sum_{k=2}^{K}\Big[ln(PC_t) \times GenRank_{kt}\Big]\eta_k \\ & + \sum_{k=1}^{k=3}\Big\{\mathbb{1}(Coll_{kt})\gamma_k + \Big[\mathbb{1}(Coll_{kt}) \times \mathbb{1}(Memb_i)\Big]\gamma_{k+3}\Big\} + \nu_t + \epsilon_{ikt} \end{aligned} \tag{1}$$

[8] The caption also conveys that the output for non-members is weighted by the ratio of the number of firms inside to outside the cartel within the entry group at any time. This weighting standardizes to a per-firm level if the number of cartel members differs from the number of non-members. It also controls for the effect of any firm exiting the generation, which can occur in the final quarters of an older chip like 16Mb. Results are similar without weighting.

The first line models the effects of accumulated output, which leads the firm to increase current output. Denote $ln(x_{ikt}) = ln\left(\sum_{\tau=1}^{t-1} q_{ik\tau}\right)$ the log of cumulative output accrued by firm i on k at $t-1$, and α_i firm fixed effects. The last term in the line interacts log cumulative output with firm fixed effects.

This specification restricts the effect of accumulated output to be equal across generations for each firm. This restriction is violated to the extent that firm learning changes over time, or firms experience significant changes to capacity levels. Its plausibility is examined by testing for pre-cartel parallel trends, and by a robustness test which considers a differential effect of cumulative output across the sample period.

The second line models the effect of cannibalization, which eventually leads the firm to reduce output; and demand, which can increase or decrease its output. Considering cannibalization, $ln(q_{irt})$ denotes industry output on the succeeding generation $r = k + 1$. The estimation of cannibalization is discussed further below.

Considering demand, define $ln(PC_t)$ as the log of worldwide PC shipments at t. The last term interacts this measure with $GenRank_{kt}$: fixed effects for the ranking of each generation from newest to oldest at t. PC shipments constituted the primary application for DRAM during the sample, and are taken as exogeneous to the supply of the component as in Igami (2017).[9] Because the marginal value of a frontier generation could be greater for new than old PCs, this specification allows an exogenous shift in demand to vary by generation ranking.

The regressor $ln(q_{irt})$ is expected to be endogenous because of unobserved factors in firm output that influence both $ln(q_{irt})$ and the dependent variable $ln(q_{ikt})$. These include temporary changes to a firm's costs such as plant outages or financial shocks in the home country. For example, a power outage at a Japanese plant forced Toshiba to cut its flash memory production by up to 20% in Q1-2010.[10]

To account for endogeneity, $ln(q_{irt})$ is instrumented by the elapsed quarters of shipment of firm i on r, denoted $FirmAge_{irt}$. $FirmAge_{irt}$ is positively correlated with $ln(q_{irt})$ over the relevant portion of r's life cycle. The condition for instrument exogeneity assumes that the initial release (the first quarter of positive output) on r is not impacted by cost shocks. This condition is more likely to hold in the semiconductor industry than others, because the predictable nature of Moore's Law permits multi-year forecasts, and each firm begins a new generation with *de minimus* production.

[9] Highlighting the importance of PC shipments, fig. C2 shows that there is little correlation between the release dates of major Windows operating systems and DRAM output, except through PC growth.

[10] See https://ca.reuters.com/article/technologyNews/idCATRE6B865620101209.

Finally, the last line estimates the effect of the cartel on output. $\mathbb{1}(Coll_{kt})$ equals one for each cartel generation during collusive periods, and $\mathbb{1}(Coll_{kt}) \times \mathbb{1}(Memb_i)$ equals one for cartel member firms in those periods. The three parameters $\{\gamma_4, \gamma_5, \gamma_6\}$ estimate the effect of the cartel on k, and ν_t denotes quarter-of-year fixed effects.

4.2 Primary Results

To begin working with the empirical model outlined above, I estimate a "pre-treatment" version of the linear regression model in equation 1. This specification treats γ as the vector of dummy variables estimating the effect of future cartel membership *prior to* the start of cartel. If the data meets the parallel trends condition of the DiD model, there should be no effect of future membership on firm output. This test can be conducted for two generations, 16Mb and 64Mb, which were active prior to the start of the cartel in 1999.

Table 2 displays the results of the regression model estimated by OLS and 2SLS. The OLS specification (1) does not include interactions between firm fixed effects and log cumulative output. Both OLS specifications estimate a negligible effect of $LogOutput_{k+1}$ on $LogOutput_k$. In contrast, 2SLS specification (3) instruments firm i's output on $k+1$ with elapsed quarters of i's output on $k+1$. Consistent with cannibalization, the estimate in (3) implies that if i doubles its output on $k+1$, it reduces its output on k by an average of about 10%. The instrument is strongly correlated with the endogenous variable (table C1 summarizes the full first stage estimates).

By accounting for cannibalization, specification (3) also estimates a higher log cumulative output effect than specification (2). A heuristic exercise can help assess its plausibility: assuming single-product static quantity-setting, symmetric firms, linear demand, and an approximately unit market elasticity, a cost reduction of 15-25% corresponds to an output increase of 30-50% with 18 firms.[11] The 44% coefficient estimate falls within this range. Similarly, observe that the effect of PC shipments varies significantly by generation rank as hypothesized. A 1% increase in PC shipments leads to a 0.48% increase in firm output on the newest generation, when production is ramping up. The magnitude of this effect rises for the second- and third-ranked generations, which are typically in the peak of their product life cycles.

[11] There were an average of 18.25 firms active across the sample period. Gardete (2016) and Siebert (2010) both estimate market-wide DRAM elasticities near -1, and Irwin and Klenow (1994) estimate DRAM learning at 20% across generations, though it varies from 10-27%.

Table 2: Estimates of Future Cartel Membership Status on Log Firm Output:
During Pre-Cartel Years

	OLS				2SLS	
	(1)		(2)		(3)	
Constant	-9.407	(2.584)	-5.035	(2.482)	-3.735	(2.643)
Log Cum Output†	0.420	(0.041)	0.366	(0.040)	0.436	(0.056)
Log Output_{k+1}‡	0.029	(0.021)	0.002	(0.019)	-0.096	(0.041)
Log PC Ship	0.816	(0.153)	0.586	(0.147)	0.481	(0.161)
Log PC Ship $\times$ Gen Rank_2	0.030	(0.010)	0.043	(0.011)	0.051	(0.012)
Log PC Ship $\times$ Gen Rank_3	-0.001	(0.013)	0.030	(0.014)	0.055	(0.017)
Log PC Ship $\times$ Gen Rank_4	-0.069	(0.013)	-0.035	(0.013)	-0.001	(0.019)
Log PC Ship $\times$ Gen Rank_5	-0.100	(0.014)	-0.072	(0.015)	-0.038	(0.018)
I(1997)_{16}	0.168	(0.193)	0.532	(0.311)	0.519	(0.324)
$\text{I(1997)}_{16} \times \text{I(Memb)}_{16}$	0.271	(0.176)	0.055	(0.350)	0.102	(0.371)
I(1998)_{16}	-0.054	(0.206)	0.174	(0.216)	0.222	(0.225)
$\text{I(1998)}_{16} \times \text{I(Memb)}_{16}$	0.413	(0.257)	0.254	(0.244)	0.289	(0.249)
Wald { 97 $\times$ I(Memb) + 98 $\times$ I(Memb) = 0 }		(0.060)		(0.570)		(0.501)
I(1997)_{64}	-0.520	(0.182)	-0.094	(0.213)	-0.010	(0.238)
$\text{I(1997)}_{64} \times \text{I(Memb)}_{64}$	0.559	(0.241)	0.127	(0.263)	-0.069	(0.261)
I(1998)_{64}	0.203	(0.256)	0.549	(0.206)	0.480	(0.262)
$\text{I(1998)}_{64} \times \text{I(Memb)}_{64}$	0.512	(0.247)	0.126	(0.222)	-0.044	(0.272)
Wald { 97 $\times$ I(Memb) + 98 $\times$ I(Memb) = 0 }		(0.003)		(0.526)		(0.802)
rk Wald F Stat*					139.76	
Firm FE $\times$ Log Cum Output	N		Y		Y	
Quarter FE	Y		Y		Y	
R-squared	0.592		0.686		0.660	
N	2,387		2,387		2,387	
Clusters	136		136		136	

Coefficients and standard errors displayed. Standard errors clustered at firm-generation level.
† Effect of log cumulative output is given for average firm.
‡ (2) instruments Log Output_{k+1} with elapsed quarters of firm i's output on $k+1$
* F statistic valid for clustering at firm-generation level, based on Kleibergen and Paap (2006).
The dependent variable is logged firm-level output of DRAM units. $\text{I(Year)}_k \times \text{I(Memb)}_k$ estimates the yearly difference in output for future cartel members relative to non-members in generation k. Sample for each regression is restricted to firm shipments within first 12 years of generation activity and at least 1000 units of quarterly shipments.

To account for the possibility of serially correlated errors, standard errors are clustered at the firm-generation level. 2SLS estimation fails to reject the null hypothesis of no pre-cartel difference in trend between members and non-members. As another specification test, I relaxed the model's restriction that the effect of cumulative output is equal across all generations for each firm. Table C2 displays the results of a subsample test with the added interaction terms Log Cum Output $\times$ Firm $\times$ $\mathbb{1}(Post)$, where $\mathbb{1}(Post) = 1$ if the generation is among the last four. For most firms in the dataset, table C2 shows small changes in the effect of cumulative output in the last four generations. This suggests that the model is appropriately specified.

Table 3 presents the results of the model that replaces pre-cartel trends with collusion indicators. In interpreting the results, it is instructive to focus on the difference between the coefficients of interest across generations. Specification (2) estimates that member output in the 16Mb generation declined significantly relative to non-members. In contrast, member outputs for the 64Mb and 128Mb generations cannot be rejected as different from non-members. Two-sided Wald tests for equality between the 16Mb member coefficient and the 64Mb and 128Mb member coefficients are rejected at 10% and 1% levels, respectively.

For several reasons, it is more difficult to estimate how effective collusion was on a single generation than across generations. The coefficients of interest have relatively large standard errors. One cause is that for each firm-generation pair, there were only 10 quarters in which the cartel was active. A second is that the magnitude of output change during the cartel is expected to vary across its members when they differ in costs that are correlated over time (e.g., Athey and Bagwell, 2008). This type of variation is consistent with fig. 5, which shows greater relative output reductions for one group of firms than another.

If firms choose output levels (as in the theoretical model below), strategic substitution implies that a decrease in output by cartel members is met with an increase by non-members. Consequently, the coefficient estimates should be interpreted such that cartel members reduced output on the 16Mb chip by some amount between $(e^{-0.61} - 1) = -46\%$ and 0%. The precise value depends on the slope of best response functions between firms, which are not recovered by the model.

By assuming that the estimated member coefficient of 46% on 16Mb consists equally of a 23% output decrease by member firms and 23% output increase by fringe firms, it is possible to construct a 95% confidence

Table 3: Estimates of Cartel Membership Status on Log Firm Output:
During Cartel Years

	OLS		2SLS	
	(1)		(2)	
Constant	-3.061	(2.207)	-2.181	(2.336)
Log Cum Output†	0.378	(0.040)	0.442	(0.057)
Log Output_{k+1}‡	0.002	(0.019)	-0.084	(0.040)
Log PC Ship	0.472	(0.132)	0.392	(0.145)
Log PC Ship × Gen Rank_2	0.044	(0.010)	0.050	(0.010)
Log PC Ship × Gen Rank_3	0.021	(0.013)	0.043	(0.015)
Log PC Ship × Gen Rank_4	-0.046	(0.012)	-0.014	(0.017)
Log PC Ship × Gen Rank_5	-0.079	(0.014)	-0.048	(0.017)
I(Collusion)_{16}	0.804	(0.125)	0.679	(0.182)
I(Collusion)_{16} × I(Memb)	-0.712	(0.225)	-0.611	(0.251)
I(Collusion)_{64}	0.441	(0.214)	0.293	(0.211)
I(Collusion)_{64} × I(Memb)	-0.213	(0.236)	-0.021	(0.223)
$\text{I(Collusion)}_{128}$	-0.231	(0.157)	-0.117	(0.213)
$\text{I(Collusion)}_{128}$ × I(Memb)	0.264	(0.196)	0.323	(0.233)
Wald { 16 × I(Memb) - 64 × I(Memb) = 0 }		(0.099)		(0.077)
Wald { 16 × I(Memb) - 128 × I(Memb) = 0 }		(0.001)		(0.005)
Firm FE × Log Cum Output	Y		Y	
Quarter FE	Y		Y	
R-squared	0.685		0.665	
N	2,387		2,387	
Clusters	136		136	

Coefficients and standard errors displayed, with clustering at firm-generation level.
† Effect of log cumulative output is given for average firm.
‡ (2) instruments Log Output_{k+1} with elapsed quarters of firm i's output on $k+1$.
The dependent variable is logged firm-level output of DRAM units. I(Collusion)_k × I(Memb) estimates the difference in output for cartel members relative to non-members during collusion in generation k. Sample for each regression is restricted to firm shipments within first 12 years of generation activity and at least 1000 units of quarterly shipments.

interval for net output change on the generation during the cartel period. Using the delta method to approximate standard errors, total 16Mb output decreased by between 69,043 to 712,584 units, or 3.1% to 24.6%, during the cartel. Allocating a larger share of the 46% estimate to members' output reduction shifts this interval higher in absolute value.[12]

4.3 Other Empirical Explanations

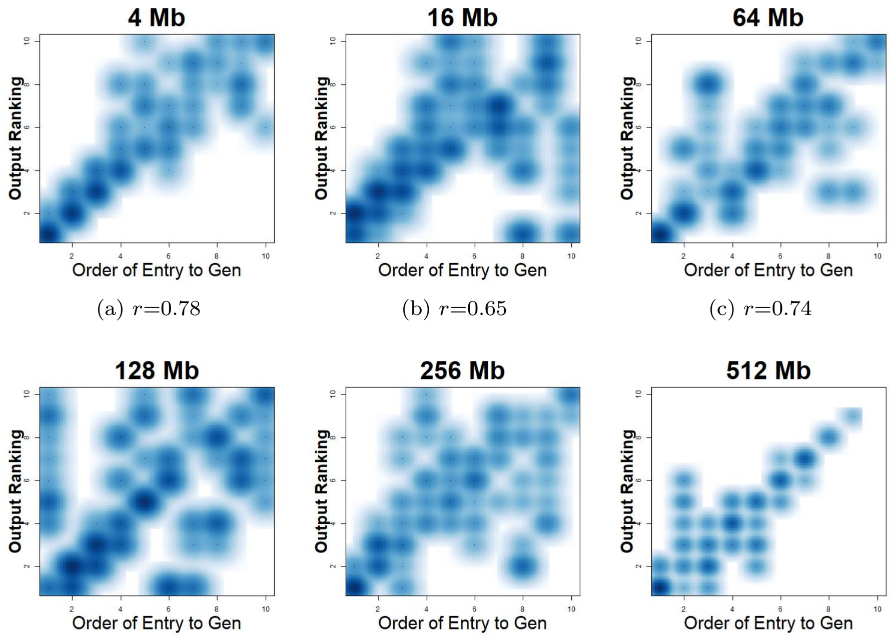
There are alternative explanations for the above patterns that do not rely upon incentive compatibility constraints. For example, cartel members in newer generations may have sought to lower costs rapidly to prevent fringe firms from establishing share. Similarly, although there is documented communication between sales representatives on 128Mb pricing, it is possible that higher level executives—who would take learning-by-doing incentives into account—did not have enough volume data to negotiate broader terms. Consistent with these scenarios, de Roos (2006) suggests that a price war in the 1990s lysine cartel could have been caused by either attempted predation or building up market share prior to a future agreement.

To address such possibilities, I construct a test for defection from collusive equilibria by one or more firms in learning-by-doing industries. Assume that firms select quantities with subsequently perfect monitoring. If an unsuccessful attempt to collude stems from IC constraints, then the episode must have featured either the *threat* of deviation or *realized* deviation. If realized, a deviating firm would expand its output relative to rivals and, if detection is not sufficiently fast, overtake rivals who had initial leads. This type of "rank reversal" is uncommon during the early years of a DRAM generation, so an increase in its level may be observed for the 64Mb and 128Mb generations.

Figure 6 plots each firm's order of entry into a generation against its subsequent output ranking during the first four years of that generation's life cycle.[13] The scatterplots are shaded to indicate the density of x-y pairs across quarters. The region of highest density tends to track a 45 degree line, reflecting a positive correlational coefficient between the two variables. This pattern is weakest for the 128Mb generation. This

[12] Using a monthly dataset with proprietary manufacturer costs, direct purchasers' economic expert estimated the average output reduction at the industry level during the collusive periods. The model yielded a 95% confidence interval of -0.8% to -13.6%. This appears to be the only estimate of the cartel's effect on output conducted by either side during civil litigation; the case was ultimately settled without verdict.

[13] For the 64Mb and 128Mb generations, the set of firms is limited to cartel members. The results are similar if all generations are limited to cartel members.



(a) r=0.78 (b) r=0.65 (c) r=0.74

(d) r=0.43 (e) r=0.70 (f) r=0.83

Figure 6: Firm Order of Entry against Subsequent Ranking, by Generation
r = sample correlation coefficient

strengthens the hypothesis that at least for the 128Mb generation, the regression results above stem from IC constraints rather than alternative explanations.

One way to formalize this test is to run the following linear regression model:

$$OutputRank_{ikt} = \beta_0 + EntryOrder_{ik}\beta_1 + \kappa_k \times \tau_t + \sum_{k=2}^{K} \Big[EntryOrder_{ik} \times \kappa_k\Big]\beta_k + \epsilon_{ikt} \tag{2}$$

The output rank and entry order are defined as above. The first interacted term denotes generation × time fixed effects, which reflect the unit of observation being studied. The coefficient on the second interacted term estimates the average correlation between entry order and output rank separately on any generations of interest k.

Table 4 below presents the results of the regression using two different samples: specifications (1) and (2) include all generations with at least 100 observations, dating to 4Kb. Specification (3) uses only those generations adjacent to the cartel: 4Mb - 512Mb. Specification (1) estimates the mean correlation for all generations. Specifications (2) and (3) estimate the mean correlation across all non-cartel generations, less the mean correlation of each of the two generations of interest.

Table 4: Output Rank Regressed on Entry Order, by Generation

	All Generations				Adjacent Generations	
	(1)		(2)		(3)	
Constant	1.402	(0.114)	1.369	(0.113)	2.045	(0.188)
Entry Order			0.802	(0.017)	0.680	(0.032)
Entry Order × I(4 **Kb**)	0.918	(0.064)				
× I(16)	0.910	(0.039)				
× I(64)	0.798	(0.060)				
× I(256)	0.932	(0.047)				
× I(1 **Mb**)	0.815	(0.049)				
× I(4)	0.716	(0.040)				
× I(16)	0.596	(0.053)				
× I(64)	0.721	(0.082)	-0.081	(0.085)	0.030	(0.100)
× I(128)	0.364	(0.073)	-0.438	(0.076)	-0.321	(0.091)
× I(256)	0.661	(0.068)				
× I(512)	0.802	(0.117)				
× I(1 **Gb**)	0.715	(0.188)				
Gen FE × Time FE	Y		Y		Y	
R-squared	0.625		0.613		0.428	
N	1,590		1,590		781	

Coefficients and standard errors displayed. Specifications (1) and (2) include all generations with at least 100 firm-time observations. (3) includes only 4Mb-512Mb generations. In all specifications, 64Mb and 128Mb generations limited to cartel members, and all generations limited to first 4 years of observations. Ties in order of entry broken based on initial output level.

In both (2) and (3), the coefficient of the 128Mb interaction is negative and significant at the 1% level. To determine whether the larger or smaller sample is more appropriate, I conduct an F-test for the joint equality of coefficients on all non-cartel generations. The null hypothesis of equality is rejected if all generations are included (as in (1)), but fails to be rejected if only 4Mb-512Mb generations are included (p-value=0.38). This suggests that the regression sample should be limited to generations adjacent to the cartel, at the cost of a smaller sample size. More distant generations are more likely to differ in the number and type of firms, which likely impacts the relevant correlation.

It is also important to test the model's sensitivity to the chosen time period for the expansion phase of a generation's life cycle. Because the test assumes that each firm is increasing its output, a longer time period increases noise if some firms begin to cut output on k to mitigate cannibalization on $k+1$. The time period above is chosen as four years, and the results are not sensitive to alternative definitions of three to five years.

This test may be revisited in future empirical studies of collusion in high technology markets. In addition to the considerations above, the test's explanatory power increases with the pace of firm learning and the frequency of data collection relative to the detection lag. It is also more informative if the identities of cheating firms are known. These firms may be named directly in the litigation record, or inferred by the researcher who matches participants' private price or share agreements to public data.[14]

5 Theoretical Model and Policy Implications

5.1 Preliminaries

The preceding empirical sections suggest that owing to incentive compatibility constraints, the DRAM cartel restricted output more significantly in the older 16Mb generation than newer generations. To expand upon these findings, this section formalizes the conditions under which it is strictly more difficult to collude on newer generations. It then discusses several implications for antitrust policy in high technology markets.

I present a two-period model of learning based on Siebert (2010) and Fudenberg and Tirole (1983). The most important feature of this model is that learning occurs only once: between a product's first and second phase. This framework permits a simple representation of dynamic effects: several years of firm learning are compressed into just one period within the model.

Consider a multi-period duopoly between two firms i and j. Each firm produces two products *old* and *new*, indexed by $k \in \{o, n\}$. Within each period, i and j set their output simultaneously a la Cournot. In the following period, each firm perfectly observes its rivals' previous output decisions. The timing of the game, which is symmetric between i and j, is depicted visually in fig. 7:

[14]One potential candidate for this type of study is the mid-2000s cartel in LCD screens. If cartel members cheated from agreements, then they may be identifiable through meetings logs which contain the cartel's target prices. These records from the so-called "crystal meetings" are included in public expert reports.

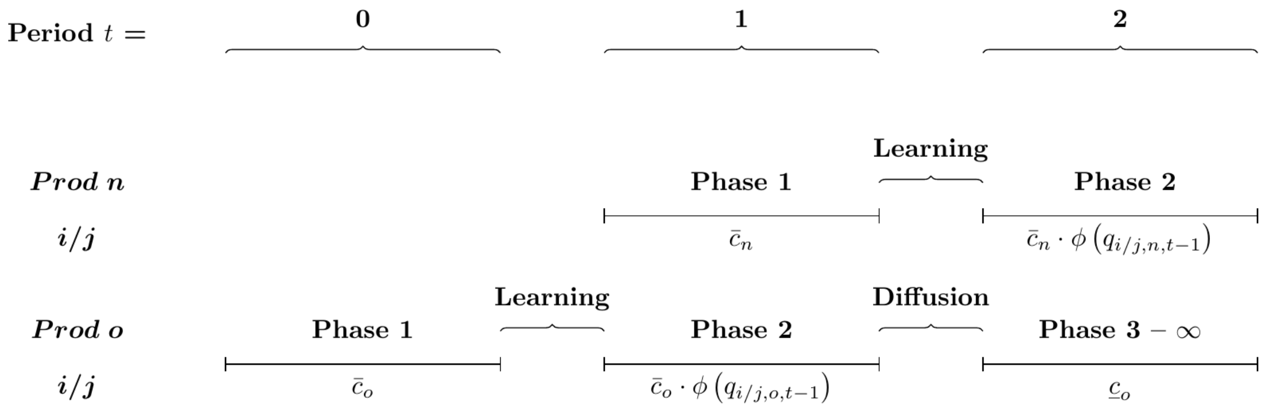


Figure 7: Timing of Product Life Cycles and Learning

There is only one supply side difference between the two products: firms begin producing n one period after o. Consequently, when firms on n engage in learning, firms on o have already finished learning. Across periods, firm learning lowers firm i or j's marginal costs from the starting point $\bar{c}_k$ to $\bar{c}_k \cdot \phi\left(q_{i/j,k,t-1}\right)$, where $\phi(0) = 1$ and $\phi(\cdot)$ is decreasing. To maintain an infinite horizon, the model adds a third phase of repeated static competition. Between phases two and three, any further learning effects on k diffuse to both firms.[15]

Next, consider the demand for each product to be given by the function $P_k(Q_{ot}, Q_{nt})$, where $Q_{kt} = \sum_{i,j} q_{i/j,kt}$. Demand is fixed across time and there is only one demand side difference between the two products: n is higher quality than o. A fraction of consumers derive greater utility from higher quality, whereas the rest are indifferent. This relation is mapped onto the inverse demand curves by assuming that there is a price intercept; that n has a higher price intercept than o; and that they have identical slopes.[16] Define the difference in price intercepts as $\mu(a) = P_n(0,0) - \underbrace{P_o(0,0)}_{a} > 0$. When $\mu(a) = 0$, all consumers are indifferent; as $\mu(a)$ increases, an increasing share prefer n. The overall size of the market increases with a.

Given a common discount factor δ, the supply and demand side of the game give rise to the following profit function for firm i on product k at time t:

$$\max_{q_{ikt}} \Pi_i = \sum_{k=o}^{n} \sum_{t=1}^{\infty} \delta^{t-1} \Big[P\left(Q_{kt}; Q_{-kt}\right) q_{ikt} - C\left(q_{ikt}, q_{ikt-1}\right) \Big] \tag{3}$$

[15] Appendix B.1 describes further conditions governing how cost enters the model.

[16] For example, a linear demand curve of this type has the form $P_{kt} = a_k - \eta Q_{kt} - \gamma Q_{-kt}$, where the intercept varies by product but the slope parameters do not.

Firms can play a variety of repeated game strategies, and the relevant equilibrium concept is subgame perfection: strategies must constitute a Nash equilibrium at every period of the game. Appendix B.1 gives the technical conditions on supply and demand sufficient for the existence of a subgame perfect equilibrium.

5.2 Collusive Equilibrium

Next, I assess the incentives of firms i and j to enforce a collusive equilibrium in the presence of learning-by-doing and multiproduct competition. Assume that firms can agree to produce the joint profit maximizing output, and to punish deviation from this output with a "grim trigger." The joint profit maximizing output is considered as a benchmark: if firms cannot reach it on a given generation, then they must resort to a sub-optimal collusive strategy or no collusive strategy at all.[17]

Formally, denote the following strategy σ'_{ik}: firms i and j maximize the discounted sum of joint profits on k beginning from period 1, and punish deviation on product k with Cournot reversion on k forever after. Equilibrium requires incentive compatibility:

Definition 1. *Define* * *and* ** *respectively as noncooperative and cooperative actions of the repeated game. To constitute a subgame perfect equilibrium,* σ'_{ik} *must satisfy:*

$$\max_{q_{ikt}, q_{jkt}} \sum_{t=1}^{\infty} \delta^{t-1} \Big[\Pi_{ikt} + \Pi_{jkt} \Big] \quad \textit{subject to} \quad \sum_{t=1}^{\infty} \delta^{t-1} \Big[\Pi^{**}_{ikt} - \Pi^{*}_{ikt} \Big] \geq 0 \tag{4}$$

Before assessing the viability of this SPE, it is useful to define the statistic $\tilde{\mu}_t(a)$. This statistic marks a convenient guide post for the extent of vertical differentiation toward n.

Definition 2. *Given* (c_{iot}, c_{int}), $\exists$ *unique* $\mu(a)$ *such that* $\Pi^*_{int} = \Pi^*_{iot}$. *Call this* $\tilde{\mu}_t(a)$.

Proof. If $\mu(a) = 0$, $\Pi^*_{iot} > \Pi^*_{int}$. As $\mu \to \infty$, $\Pi^*_{int} > \Pi^*_{iot} \to 0$. By continuity of profits, there exists an intersection point. By decreasing differences, $\Pi^*_{int} - \Pi^*_{iot}$ is increasing in μ, so the intersection is unique. $\square$

[17] Other types of trigger and stick-and-carrot strategies share the property that the extent of punishment is based on marginal cost. By focusing on the relationship between learning and marginal cost, the model below highlights the intuition for a variety of collusive strategies.

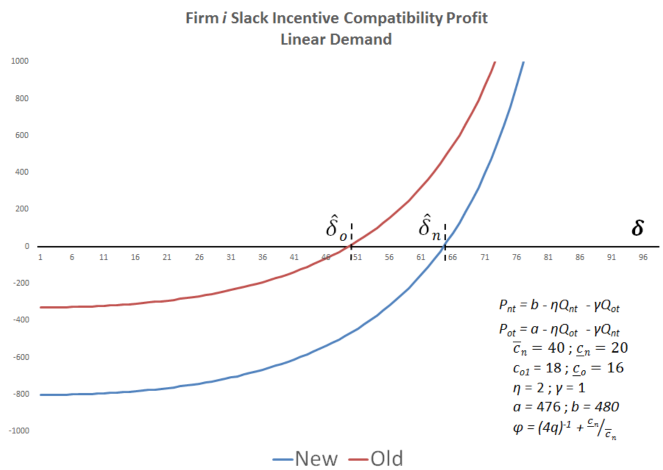


Figure 8: Minimum Discount Factor, n and o

Definition 2 shows that under competitive equilibrium, in each period there is a unique level of vertical differentiation such that the profits on each generation are equal. This occurs because firms produce n at a cost disadvantage relative to o in each period. Given this definition, we are ready to prove the main result:

Proposition 1. *Define $\hat{\delta}_k$ as the minimum discount factor that satisfies the ICC for σ'_{ik}.*
If $\mu(a) \leq \tilde{\mu}_1(a)$, then $\hat{\delta}_o < \hat{\delta}_n$.

Proof. See Appendix B.2. □

Proposition 1 formalizes the intuition that it is strictly more difficult to collude on a new generation than an older generation if the learning rate decreases monotonically over time. The sufficient condition is that vertical differentiation is small enough that profits on n are no greater than profits on o at the start of collusion. This condition assumes symmetric firms and is sufficient but not necessary.[18]

Figure 8 depicts the result graphically with a set of parameters that satisfy the condition. As δ increases, so does the profit on the incentive constraint, eventually making σ'_{ik} sustainable. Because slack on o lies strictly above n, $\hat{\delta}_o < \hat{\delta}_n$. When $\hat{\delta}_o < \delta < \hat{\delta}_n$, σ'_{ik} is sustainable for o but not n.

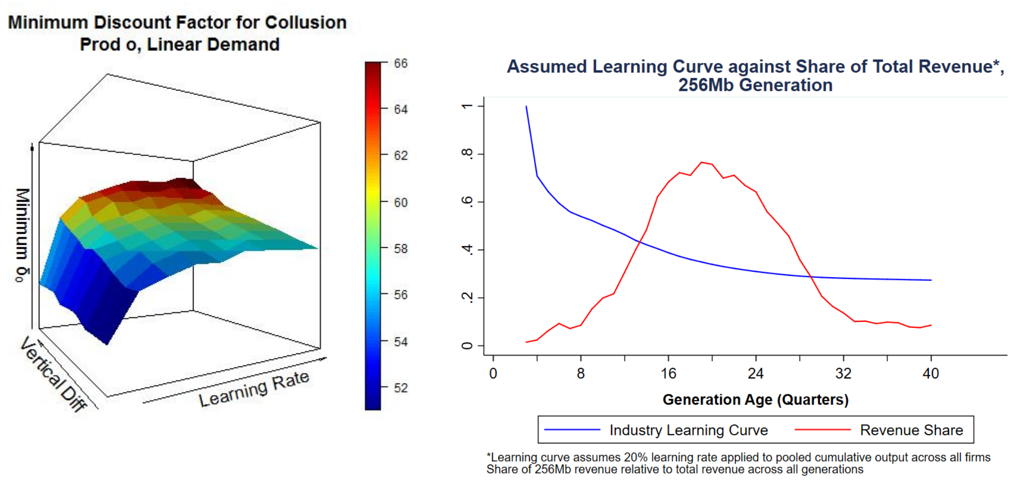
Proposition 1 also assumes that firms formulate strategies that treat different products independently. Appendix A permits firms to punish deviation on n with conduct on o, and shows that joint collusion on n and o remains strictly more difficult than collusion on o.

[18] At the start of the DRAM cartel in Q2-1999, the sufficient condition was partially satisfied: $\Pi_{128} < \Pi_{16} < \Pi_{64}$.

5.3 Discussion

The modeling framework can be used to highlight wider issues relevant to antitrust policy. In a report to the European Commission, Ivaldi et al. (2003) present a model of collusion in innovative markets. Their model shows that as the exogenous probability that any firm achieves a future innovation increases, collusion between duopolists is more difficult to sustain.

Many high technology markets are characterized by innovation through learning-by-doing as well as overlapping product generations.[19] Innovation through learning-by-doing takes place endogenously through a firm's output. The effect of innovation on collusion can therefore be separated from the effect of future demand on collusion; competition agencies may find both forces important in assessing coordinated effects in high technology mergers.



(a) $P_{ot} = a - 2Q_{ot} - Q_{nt}$; $P_{nt} = \boldsymbol{b} - 2Q_{nt} - Q_{ot}$
$a = 476$; $\boldsymbol{b} \in [476, 521]$
$\phi(q) = (\boldsymbol{l}q)^{-1} + \underline{c}_n/\bar{c}_n$; $\boldsymbol{l} \in [1, 10]$

(b)

Figure 9

To illustrate, fig. 9 plots a theoretical numerical simulation (panel a) alongside empirical data from the 256Mb generation (panel b). Panel (a) plots the minimum discount factor necessary to sustain collusion on generation o during its learning phase. By inspection, the critical value increases with the pace of learning and the degree of vertical differentiation toward n. Moreover, the effect of increased learning on $\hat{\delta}_o$ intensifies

[19]Prominent examples include microprocessors, servers, LCD panels, hard drives, and numerous other computer hardware and electronics products.

as vertical differentiation is higher. The reason is that if n is significantly more attractive than o, a reduction in n's cost speeds up the the obsolescence of o.

Panel (b) plots a simplified representation of learning-by-doing against the total revenue share of the 256Mb generation, the first after the discovery of the cartel. Learning is modeled as $C_{kt} = aX_{kt}^{-b}$, where $X_{kt} = \sum_i x_{ikt}$ is the cumulative output accrued by the industry on generation k at time $(t-1)$; $b = -0.2$ indicates a 20% learning rate; and $a = 1$ is the initial cost.[20] The graph shows that the revenue share of the 256Mb generation is highest when most learning has likely been completed, and remains significant for several more quarters. This suggests that vertical differentiation is sufficiently low that DRAM generations are long-lived, and consequently, that collusion may pose a threat after several years of production. Agencies can construct this type of measure using commonly available data: learning rates may appear in firms' business documents, proprietary engineering estimates, or the academic literature. Similarly, the speed of obsolescence can be gauged by calculating revenues or conducting customer surveys.

The effects of learning-by-doing on collusion should also be considered separately from other types of innovation. It is well understood that the limits of nanoscale technology are beginning to inhibit semiconductors from achieving the rapid cost gains of the preceding decades. On the other hand, the same phenomenon is spurring investment in R&D geared toward new chip technologies.[21] The former trend should increase the viability of collusion whereas the latter should decrease it (as in Ivaldi et al., 2003). In evaluating coordinated effects, the competition agency can investigate managers' projections of future learning rates, the most likely future innovators, and the stage of new product development. It is possible that this evaluation would suggest a larger threat of coordination in the short- versus long-term, which the agency should weigh in light of the appropriate competition law environment.

6 Conclusion

If firms in innovative industries successfully refrain from competition, the cost to society could be especially large, because firms would also delay the rate at which they reduce long-term costs. The fundamental insight of this paper is that the effectiveness of collusion in such markets is determined by the rate of learning. In

[20] As footnote 11 notes, Irwin and Klenow (1994) estimate DRAM learning to be 20% on average across generations.

[21] For example, "Computer Chips Evolve to Keep Up with Deep Learning," *Wall Street Journal* 1/11/2017.

particular, when firms simultaneously produce multiple products, it can be more difficult to sustain collusion on the newer product(s). Empirical analysis of the DRAM cartel is consistent with this scenario.

More generally, this paper illustrates that learning-by-doing is one characteristic through which cartels in a dynamic setting may face different challenges than those in static ones. This possibility depends on the institutional elements of each market, including firm patience, the pace of learning, the number of products each firm sells, and the extent of vertical differentiation between them. Nonetheless, several other known collusive episodes in high-technology markets raise fruitful opportunities for further research. Careful study of those cases may reveal more about the challenges and strategies taken by modern cartels.

References

Athey, Susan, and Kyle Bagwell. 2008. "Collusion with persistent cost shocks." *Econometrica*, 76(3): 493–540.

Bernheim, Douglas, and Michael D. Whinston. 1990. "Multimarket Contact and Collusive Behavior." *Rand Journal of Economics*, 21(1): 1–26.

Cabral, Luis, and Michael Riordan. 1997. "The Learning Curve, Predation, Antitrust, and Welfare." *The Journal of Industrial Economics*, 45(2): 155–169.

Clark, Robert, and Jean-Francois Houde. 2013. "Collusion with Asymmetric Retailers: Evidence from a Gasoline Price-Fixing Case." *American Economic Journal: Microeconomics*, 5(3): 97–123.

Clark, Robert, and Jean-Francois Houde. 2014. "The Effect of Explicit Communication on Pricing: Evidence from the Collapse of a Gasoline Cartel." *The Journal of Industrial Economics*, 62(2): 191–228.

Dasgupta, Partha, and Joseph Stiglitz. 1988. "Learning-by-Doing, Market Structure and Industrial and Trade Policies." *Oxford Economics Papers*, 40(2): 246–268.

de Roos, Nicolas. 2004. "A Model of Collusion Timing." *International Journal of Industrial Organization*, 22(3): 351–387.

de Roos, Nicolas. 2006. "Examining models of collusion: The market for lysine." *International Journal of Industrial Organization*, 24(6): 1083–1107.

Ellison, Glenn. 1994. "Theories of Cartel Stability and the Joint Executive Committee." *The Rand Journal of Economics*, 25(1): 37–57.

Fudenberg, Drew, and Jean Tirole. 1983. "Learning-by-Doing and Market Performance." *The Bell Journal of Economics*, 14(2): 522–530.

Gardete, Pedro. 2016. "Competing Under Asymmetric Information: The Case of Dynamic Random Access Memory Manufacturing." *Management Science*, 62(11): 3291–3309.

Genesove, David, and Wallace P. Mullin. 2001. "Rules, Communication and Collusion: Narrative Evidence from the Sugar Institute Case." *American Economic Review*, 91(3): 379–398.

Ghemawat, Pankaj, and Michael A. Spence. 1985. "Learning Curve Spillovers and Market Performance." *The Quarterly Journal of Economics*, 100: 839–852.

Igami, Mitsuru. 2017. "Estimating the Innovator's Dilemma: Structural Analysis of Creative Destruction in the Hard Disk Drive Industry, 1981-1998." *Journal of Political Economy*, 125(3): 798–847.

Igami, Mitsuru, and Takuo Sugaya. 2017. "Measuring the Incentive to Collude: The Vitamin Cartels, 1990-1999." Working Paper, Yale University.

Irwin, Douglas A., and Peter J. Klenow. 1994. "Learning-by-Doing and Spillovers in the Semiconductor Industry." *Journal of Political Economy*, 102(6): 1200–1227.

Ivaldi, Marc, Bruno Jullien, Patrick Rey, Paul Seabright, and Jean Tirole. 2003. "The Economics of Tacit Collusion." Final Report for DG Competition, European Commission.

Johnson, Justin P., and David P. Myatt. 2006. "Multiproduct Cournot Oligopoly." *The Rand Journal of Economics*, 37(3): 583–601.

Kleibergen, Frank, and Richard Paap. 2006. "Generalized reduced rank tests using the singular value decomposition." *Journal of Econometrics*, 133(1): 97–126.

Levenstein, Margaret. 1997. "Price Wars and the Stability of Collusion: A Study of the Pre-World War I Bromine Industry." *The Journal of Industrial Economics*, 45(2): 117–137.

Levenstein, Margaret, and Valerie Suslow. 2006. "What Determines Cartel Success?" *Journal of Economic Literature*, 44(1): 43–95.

Mariuzzo, Franco, and Patrick Paul Walsh. 2013. "Commodity Market Dynamics and the Joint Executive Committee, 1880-1886." *Review of Economics and Statistics*, 95(5): 1722–1739.

Mookherjee, Dilip, and Debraj Ray. 1991. "Collusive Market Structure under Learning-by-Doing and Increasing Returns." *Review of Economics Studies*, 58(5): 993–1009.

Noll, Roger. 2014. "The DRAM Antitrust Litigation." In *The Antitrust Revolution: Economics, Competition and Policy.* . One ed., , ed. John E. Kwoka Jr. and Lawrence J. White, 246–276. Oxford University Press.

Porter, Robert. 1983. "A Study of Cartel Stability: The Joint Executive Committee, 1880-1886." *The Bell Journal of Economics*, 14(2): 301–314.

Roller, Lars-Hendrik, and Frode Steen. 2006. "On the Workings of a Cartel: Evidence from the Norwegian Cement Industry." *The American Economic Review*, 96(1): 321–338.

Scott Morton, Fiona. 1997. "Entry and Predation: British Shipping Cartels 1879-1929." *Journal of Economics and Management Strategy*, 6(4): 679–724.

Siebert, Ralph. 2010. "Learning-by-Doing and Cannibalization Effects at Multi-Vintage Firms: Evidence from the Semiconductor Industry." *The B.E. Journal of Economic Analysis & Policy*, 10(1): Article 37.

Spence, Michael A. 1981. "The Learning Curve and Competition." *The Bell Journal of Economics*, 12(1): 49–70.

Appendices

A Collusion with Multiproduct Punishments

The discussion in the body permitted only one type of collusive strategy: $\sigma_{ik}^{'}$. Now, permit a second collusive strategy $\sigma_i^{''}$. This strategy also requires each firm to maximize the discounted sum of joint profits, but punishes deviation on generation k with Cournot reversion on k *and* $-k$. When punishments occur across products, firm i will consider either adhering to the collusive path on both generations or deviating from the collusive path on both generations.[22]

For subgame perfection, firms must satisfy the following condition:

$$\max_{q_{ikt}, q_{jkt}} \Pi_i + \Pi_j \quad \textit{subject to} \quad \sum_{k=o}^{n} \sum_{t=1}^{\infty} \delta^{t-1} \left[\Pi_{ikt}^{**} - \Pi_{ikt}^{*} \right] \geq 0 \tag{5}$$

Equation (5) differs from the single-product case, eq. (4), in only one way: firms can pool IC constraints between products (Bernheim and Whinston, 1990). Define the IC constraints corresponding to o and n as $g_o(\delta)$ and $g_n(\delta)$, respectively. The constraint in eq. (5) can be decomposed as:

$$g_o(\delta) + g_n(\delta) \geq 0 \tag{6}$$

$$g_o(\delta) \geq -g_n(\delta) \tag{7}$$

Equation (6) shows that $\sigma_i^{''}$ is sustainable only if the *sum* of the individual IC constraints is greater than or equal to zero. If $\delta > \hat{\delta}_o$, then $g_o(\delta) > 0$: there is "complementary slack" on o. Equation (7) shows that this slack can be redistributed: if $g_o(\delta)$ is sufficiently large relative to (negative) slack $-g_n(\delta)$, then collusion can be sustained on both generations. This logic leads to the following result:

Proposition A1. *Define $\hat{\hat{\delta}}$ as the minimum discount factor that satisfies the ICC for σ_i''.*

If $\mu(a) \leq \tilde{\mu}_1(a)$, then $\hat{\delta}_o < \hat{\hat{\delta}} < \hat{\delta}_n$.

Proof. See Appendix B.3. ☐

[22] Deviation on k and adherence on $-k$ is strictly dominated by deviation on both.

Proposition A1 imposes the same vertical differentiation condition as Proposition 1, and carries two implications. First, collusion on both products under $\sigma_i^{''}$ is *more* difficult to sustain than collusion on product o alone. The reason is simple: to generate any improvement relative to the single-product case, the multiproduct case requires at least some slack profit on o, which in turn requires firms to be more patient. Second, and for the same reason, collusion on both products is *more* sustainable with multiproduct than single-product punishment.

In sum, the effect of Proposition A1 is that it "narrows the gap" between minimum discount thresholds demonstrated in Proposition 1. However, whereas the gap is reduced, it cannot be completely eliminated. For $\delta \in (\hat{\delta}_o, \hat{\hat{\delta}})$, σ_i'' and σ_{in}' are unsustainable, but σ_{io}' remains sustainable.

B Proofs of Results

B.1 Technical Conditions Guaranteeing Equilibrium

On the supply side, each firm's marginal costs evolve based on $c_k \in [\underline{c}_k, \bar{c}_k]$. At phase one, firms have equal marginal cost $\bar{c}_k$ within a generation, but asymmetric costs across generations: $\bar{c}_o < \bar{c}_n$. Between phases one and two, marginal cost declines from $\bar{c}_k$ to $\bar{c}_k \cdot \phi(q_{ikt-1})$. There are diminishing returns to learning, such that $\phi(0) = 1$, $\frac{\partial \phi}{\partial q} < 0$, $\frac{\partial^2 \phi}{\partial q^2} > 0$. Marginal cost is bounded from below such that $\lim_{q \to \infty} \phi(q) = \underline{c}_k / \bar{c}_k$. At phase three, marginal cost reaches its minimum $\underline{c}_k$, $\underline{c}_o < \underline{c}_n$.

On the demand side, define $\mathbf{P}(Q_{ot}, Q_{nt})$, $\mathbf{P} : \Re^2_+ \mapsto \Re^2_+$, where $Q_{kt} = \sum_{i,j} q_{i/j,kt}$. Assume that $\mathbf{P}$ is invertible and twice continuously differentiable. Each element $P_k \in [0, \bar{P}_k]$ with $\frac{\partial P_k}{\partial Q_k} < 0$ and $\frac{\partial P_k}{\partial Q_{-k}} < 0$ $\forall\ Q_k > 0$: the two products are substitutes. Let the inverse demand slopes be symmetric across products and fixed across time.

With constant marginal costs, restrictions on the demand curve are sufficient to guarantee equilibrium. Assume the multiproduct version of the downward-sloping marginal revenue assumption. That is, define $\mathrm{MR}_{ik} = P_k + \frac{\partial P_k}{\partial Q_k} q_{ik} + \frac{\partial P_{-k}}{\partial Q_k} q_{i-k}$ and assume:

$$\frac{\partial \mathrm{MR}_{ik}}{\partial q_{ik}} \leq 0 \text{ and } \frac{\partial \mathrm{MR}_{ik}}{\partial q_{i-k}} \leq 0 \tag{8}$$

Marginal revenue on k is weakly decreasing in k as well as $-k$. The first inequality implies that that profits display decreasing differences on k across firms, and the second that profits display decreasing differences across products k and $-k$ (within or across firms). With two firms, condition 8 guarantees the existence and uniqueness of a Cournot-Nash equilibrium at each period (Johnson and Myatt, 2006). A sufficient condition for 8 to hold is log-concave inverse demand, e.g. with the exponential form.

B.2 Proof of Proposition 1

Separate the ICC for collusion on generations $k \in \{o, n\}$ into two parts, as follows:

$$\sum_{t=2}^{\infty} \delta^{t-1} \left(\Pi_{ikt}^{**} - \Pi_{ikt}^{*}\right) \geq \Pi_{ik1}^{*} - \Pi_{ik1}^{**}$$

By construction, the left hand side is increasing in δ, and it follows that there is a critical discount factor $\hat{\delta}_k$ such that σ'_{ik} is sustainable for $\delta \geq \hat{\delta}_k$ and unsustainable for $\delta < \hat{\delta}_k$.

The sufficient condition is that $\mu(a) \leq \tilde{\mu}_1(a)$. Assume first that this constraint is binding. Further, assume the absence of any learning effects on n. Both of these assumptions will be relaxed in turn.

In assessing the incentive to deviate, we must consider two different scenarios: deviation on n, and deviation on o. Through symmetry of demand, the respective on- and off-equilibrium path profits under σ'_{ik} are equal in both scenarios, which implies equality of deviation profits:

$$\begin{aligned} \Pi_{iot}^{**} &= \Pi_{int}^{**} \\ \Pi_{iot}^{*} &= \Pi_{int}^{*} \end{aligned} \tag{9}$$

$$\sum_{t=2}^{\infty} \delta^{t-1} \left[\Pi_{int}^{**} - \Pi_{int}^{*}\right] = \sum_{t=2}^{\infty} \delta^{t-1} \left[\Pi_{iot}^{**} - \Pi_{iot}^{*}\right] \tag{10}$$

Next, consider any period t in which the sufficient condition strictly holds: $\mu(a) < \tilde{\mu}_t(a)$. By definition, $\Pi_{iot}^{*} > \Pi_{int}^{*}$. Then the marginal profitability of reducing output is greater on o than n:

$$\frac{\partial \pi_{iot}}{\partial q_{iot}} > \frac{\partial \pi_{int}}{\partial q_{int}} \tag{11}$$

It follows that total profitability in reducing output from the Cournot-Nash equilibrium to the collusive equilibrium is greater on o than n:

$$\int_{q_{iot}^{**}}^{q_{iot}^{*}} \frac{\partial \pi_{iot}}{\partial q_{iot}}\, dq > \int_{q_{int}^{**}}^{q_{int}^{*}} \frac{\partial \pi_{int}}{\partial q_{int}}\, dq \tag{12}$$

$$\Pi_{iot}^{**} - \Pi_{iot}^{*} > \Pi_{int}^{**} - \Pi_{int}^{*} \tag{13}$$

Finally, reintroduce the learning effect for product n. Consider period 2 and assume for the sake of contradiction that:

$$(\Pi_{io2}^{**} - \Pi_{io2}^{*}) \leq (\Pi_{in2}^{**} - \Pi_{in2}^{*}) \tag{14}$$

Inequality 14 implies:

$$\begin{aligned} &(P_{o2}^{**}(\cdot) - \underline{c}_o) \cdot q_{io2}^{**} - (P_{o2}^{*}(\cdot) - \underline{c}_o) \cdot q_{io2}^{*} \quad \leq \\ &(P_{n2}^{**}(\cdot) - c_{in2}\,(q_{in1}^{**})) \cdot q_{in2}^{**} - (P_{n2}^{*}(\cdot) - c_{in2}\,(q_{in1}^{*})) \cdot q_{in2}^{*} \end{aligned} \tag{15}$$

Inequality 15 implies that $c_{in2}\,(q_{in1}^{**}) \leq c_{in2}\,(q_{in1}^{*})$. However, $q_{in1}^{**} < q_{in1}^{*}$ implies:

$$\begin{aligned} c_1 \cdot \phi\,(q_{in1}^{**}) &> c_1 \cdot \phi\,(q_{in1}^{*}) \\ c_{in2}\,(q_{in1}^{**}) &> c_{in2}\,(q_{in1}^{*}) \text{※} \end{aligned} \tag{16}$$

It follows that slack at period 2 is greater on o than n:

$$\Pi_{io2}^{**} - \Pi_{io2}^{*} > \Pi_{in2}^{**} - \Pi_{in2}^{*} \tag{17}$$

Combining 10, 13, and 17, it follows that the minimum discount factor necessary to sustain σ_{io}' is less than that to sustain σ_{in}' :

$$\hat{\delta}_o < \hat{\delta}_n \quad \blacksquare \tag{18}$$

B.3 Proof of Proposition A.1

Appendix B.2 proves that $\hat{\delta}_n > \hat{\delta}_o$. Also recall from Appendix A that the pooled ICC corresponding to $\sigma_i^{''}$ is:

$$g_o(\delta) \geq -g_n(\delta) \tag{19}$$

The proof is divided into two parts to correspond to two distinct cases. First, I show that $\hat{\hat{\delta}}$ is strictly greater than $\hat{\delta}_o$. Consider increasing δ from the left. As $\delta \to \hat{\delta}_o$ from the left, $g_o(\delta) \to 0$ from the left and 19 is unsatisfied:

$$\lim_{\delta \to \hat{\delta}_o} g_o(\delta) = 0 \tag{20}$$

$$\lim_{\delta \to \hat{\delta}_o} g_o(\delta) + g_n(\delta) < 0 \tag{21}$$

$$\lim_{\delta \to \hat{\delta}_o} g_o(\delta) < -g_n(\delta) \tag{22}$$

And it follows that $\hat{\hat{\delta}} > \hat{\delta}_o$.

Second, I show that $\hat{\hat{\delta}}$ is strictly less than $\hat{\delta}_n$. Define $g_o(\hat{\delta}_n) - g_o(\hat{\delta}_o) = \gamma > 0$. Consider the range $\hat{\delta}_o < \delta < \hat{\delta}_n$. As $\delta \to \hat{\delta}_n$ from the left, $g_n(\delta) \to 0$ from the left and 19 is satisfied:

$$\lim_{\delta \to \hat{\delta}_n} g_n(\delta) = 0 \tag{23}$$

$$\lim_{\delta \to \hat{\delta}_n} g_o(\delta) + g_n(\delta) = \gamma + 0 \tag{24}$$

$$\lim_{\delta \to \hat{\delta}_n} g_o(\delta) - \gamma = -g_n(\delta) \tag{25}$$

Define $\epsilon > 0$ as an arbitrarily small constant. By the continuity of $g_o(\delta)$, there exists an ϵ such that:

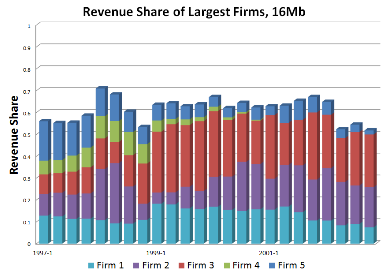
$$g_o(\hat{\delta}_n) - g_o(\hat{\delta}_n - \epsilon) \leq \gamma \tag{26}$$

$$g_o(\hat{\delta}_n - \epsilon) \geq g_o(\hat{\delta}_n) - \gamma = \underbrace{-g_n(\hat{\delta}_n)}_{=0} > -g_n(\hat{\delta}_n - \epsilon) \tag{27}$$

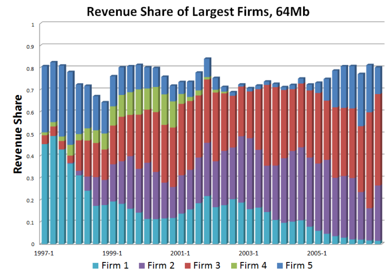
$$g_o(\hat{\delta}_n - \epsilon) > -g_n(\hat{\delta}_n - \epsilon) \tag{28}$$

And it follows that $\hat{\hat{\delta}} < \hat{\delta}_n$. ■

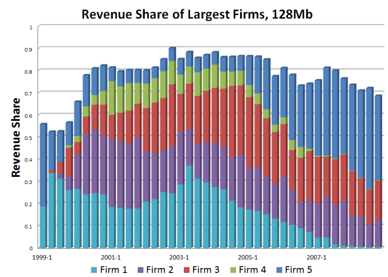
C Additional Figures and Tables



(a)



(b)



(c)

Figure C1: Market Share by Generation, Largest Firms
Firm names are masked to avoid conveying the impression of private information about cartel participants.

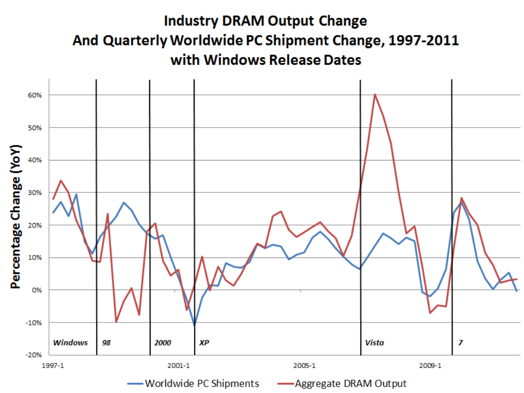


Figure C2: Industry DRAM Output against PC Shipment Change

Table C1: First Stage Estimate of Log Output$_{k+1}$ on Exogenous Variables

	OLS	
Constant	17.349	(5.793)
Firm Age$_{k+1}$‡	0.194	(0.016)
Log Cum Output†	0.510	(0.075)
Log PC Ship	-1.207	(0.329)
Log PC Ship $\times$ Gen Rank$_2$	0.069	(0.020)
Log PC Ship $\times$ Gen Rank$_3$	0.190	(0.029)
Log PC Ship $\times$ Gen Rank$_4$	0.213	(0.031)
Log PC Ship $\times$ Gen Rank$_5$	0.121	(0.031)
I(1997)$_{16}$	0.041	(0.692)
I(1997)$_{16}$ $\times$ I(Memb)$_{16}$	0.417	(0.903)
I(1998)$_{16}$	0.867	(1.064)
I(1998)$_{16}$ $\times$ I(Memb)$_{16}$	0.270	1.265
I(1997)$_{64}$	0.553	(0.571)
I(1997)$_{64}$ $\times$ I(Memb)$_{64}$	-1.621	(0.611)
I(1998)$_{64}$	-0.654	(0.573)
I(1998)$_{64}$ $\times$ I(Memb)$_{64}$	-1.429	(0.785)
Firm FE $\times$ Log Cum Output	Y	
Quarter FE	Y	
R-squared	0.754	
N	2,387	
Clusters	136	

Coefficients and standard errors displayed. Standard errors clustered at firm-generation level.
‡ Instrumental variable equals elapsed quarters of firm i producing $k+1$
† Effect of log cumulative output is given for average firm.
The dependent variable is logged firm-level output of units on generation $k+1$. Sample for each regression is restricted to firm shipments within first 12 years of generation activity and at least 1000 units of quarterly shipments.

Table C2: Firm Log Cumulative Output on Log Current Output:
Subsample Test - 2SLS Model, 1997-2011

Firm	$\times$ Log Cum Output[†]		$\times$ Log Cum Output $\times$ I(Post)[‡]	
2	-0.028	(0.033)	0.021	(0.033)
3	0.044	(0.035)	-0.029	(0.013)
4	0.038	(0.031)	0.012	(0.014)
5	-0.026	(0.032)	0.035	(0.031)
6	0.032	(0.031)	-0.021	(0.015)
7	-0.067	(0.031)	-0.041	(0.021)
8	-0.010	(0.031)	0.008	(0.018)
9	-0.084	(0.034)	0.012	(0.022)
10	-0.040	(0.030)	0.034	(0.020)
11	-0.017	(0.035)	0.026	(0.025)
12	-0.002	(0.032)	0.043	(0.018)
13	-0.070	(0.033)	-0.013	(0.023)
14	-0.002	(0.033)	-0.248	(0.035)
15	-0.016	(0.033)	-0.024	(0.019)
Log Cum Output	0.454 (0.059)			
Log PC Ship	Y			
Log PC Ship $\times$ Gen Rank$_k$	Y			
Log Output$_{k+1}$	Y			
Quarter FE	Y			
R-squared	0.644			
N	2,376			
Clusters	131			

Coefficients and standard errors (clustered at firm-generation level) displayed.
[†] estimates firm i learning effect for 1Mb - 128Mb
[‡] estimates incremental firm i learning learning effect for 256Mb - 2Gb
The dependent variable is logged firm-level output of DRAM units. Regression estimated via 2SLS by instrumenting Log Output$_{k+1}$ by Firm Age$_{k+1}$. Excludes firm-generation pairs with less than or equal to 4 quarters of data. Firm names are masked to avoid conveying the impression of private information about cartel participants.

D Trial Transcripts

Q. NOW, THERE WERE -- I WANT TO ACTUALLY TURN TO THE CHART YOU GAVE US YESTERDAY.
A. YES.
Q. WITH THE, I THINK, THE ONE SAYS DRAM AVERAGE SALES PRICE INDEX.
A. YES. I'M THERE.
Q. AND THAT CHART SHOWING IN 2001 THERE WAS A TIME PERIOD OF SIGNIFICANT PRICE INCREASE, DOESN'T IT?
A. IN 2001 FOR, LET'S SEE. RIGHT, AT THE END OF 2001 FOR THREE CHIPS THERE WAS A SIGNIFICANT INCREASE.
Q. THE 16 MEG PRICE WENT, I KNEW IT'S A ROUGH APPROXIMATION, FROM ABOUT A MEASURE OF .25 LOOKS LIKE ON YOUR LEFT SOMEWHERE BELOW .75?
A. WELL, IT WENT FROM ABOUT, YES, JUST TO BE ACCURATE, ABOUT .3 TO ABOUT .6.
Q. SO THAT A DOUBLING OF PRICE DURING THAT PERIOD?
A. APPROXIMATELY, YES.
Q. WHAT WAS THE TIME PERIOD IN WHICH THE PRICE DOUBLED?
A. THAT WOULD BE, ACCORDING TO THIS, WOULD BE THE FOURTH QUARTER OF '01 AND THE FIRST QUARTER OF, EXCUSE ME, AND THE FIRST QUARTER OF '02.
Q. THE PRICES FOR 64 AND 128 SIZE WENT UP DURING THAT SAME TIME PERIOD, BUT NOT AS MUCH?
A. THAT'S WHAT I TESTIFIED TO.

Figure D1: Excerpt from Transcript of Proceedings as to Criminal Trial of Gary Swanson, Case No. 06-0692, U.S. District Court N.D. Cal., 12/17/2007, pg. 97.