Measuring the Incentive to Collude: The Vitamin Cartels, 1990–1999*

Mitsuru Igami[†] Takuo Sugaya[‡]

July 12, 2017

Abstract

Why do some cartels survive for a decade whereas others collapse within a few years? Models of collusion are difficult to identify, but the vitamins case, one of the largest in history, entails direct evidence from American courts and European agencies. We provide a theory-based measurement of the incentive to collude, and test a fundamental prediction of game theory that cooperation is self-enforcing if and only if it is incentive compatible. Our simple repeated-game model could explain the life and death of various vitamin cartels. Simulations suggest a hypothetical merger could have prolonged the vitamin C cartel.

Keywords: Antitrust, Cartel, Collusion, Coordinated effect, Merger, Repeated game.

JEL classifications: D43, L13, L41.

*We thank John Connor and Leslie Marx for sharing institutional knowledge, Bryant Johnson for help in obtaining the U.S. court documents, and Chise Ayuha for digitizing data from these documents. We benefited from comments at MIT, Keio, Stanford, Yale, Georgetown, Cornell, and the 2017 IIOC meeting.

†Yale Department of Economics. E-mail: mitsuru.igami@yale.edu.

‡Stanford Graduate School of Business. Email: tsugaya@stanford.edu.

1 Introduction

One of the key lessons from game theory is that cooperation is self-enforcing if and only if it is incentive compatible. We test this central prediction by studying the international cartels of vitamin manufacturers, which were among the largest cases in history. The validity of this prediction has direct policy implications. We need not worry about regulation when cartels are unstable. When they are not, however, the authority would need to intervene.

Measuring the incentive to collude and its determinants is the first step toward understanding cartels and improving antitrust policy. This seemingly simple task has turned out to be challenging. On the one hand, theory says almost any outcome can be supported as an equilibrium, thanks to folk theorem,[1] which means we cannot make sharp predictions unless we know the specifics of firms' payoffs, strategies, and beliefs. Hence, we need data and institutional knowledge to make progress. On the other hand, data do not exist, because explicit collusion (i.e., cartel) is *per se* illegal. The catch is that successful cartels leave no record, and tacit collusion is tacit. Hence, seminal works such as Porter (1983) and Ellison (1994) had to study a 19th-century case before the birth of antitrust law (i.e., when the cartel was perfectly legal), which may or may not be informative about cartels of our time.

The vitamins case broke the world record of antitrust fines in 1999 and is still among the largest to date.[2] Roche, a Swiss drug company, cooperated with 20 other vitamin makers around the world and cartelized 16 different product categories in the 1990s. Individual vitamins constitute separate markets because each vitamin is used for its specific biochemical function, and production of each vitamin requires its own specific plant (UKCC 2001, p.108). Some cartels broke up within a few years; others survived until the American government busted them. Thus, the case is a laboratory of cartel stability. Detailed evidence was generated by criminal investigations by the Federal Bureau of Investigation (FBI), the Department of Justice (DOJ), and the European Commission's Directorate General of Competition (DG-Comp). More than a dozen famous economists wrote expert witness reports for class-action litigation in the early 2000s, including Dr. Douglas Bernheim and Dr. Daniel McFadden, some of whom went on to publish influential books on the economics of collusion.[3]

With this level of high-quality documentation, we can actually estimate the cartel members' payoffs and directly observe their communications, strategies, and even beliefs. Hence, the vitamins case represents both a major episode in the history of antitrust enforcement

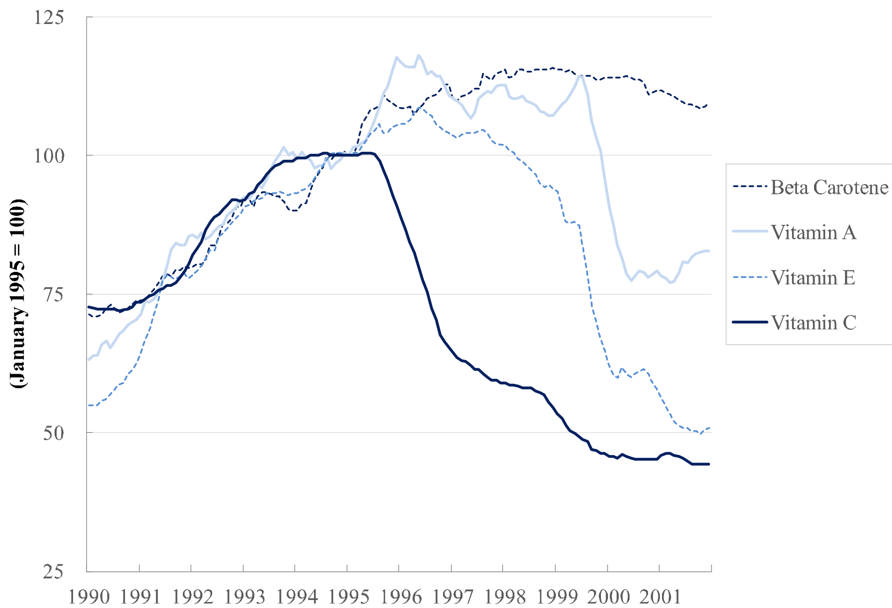
[1] See Fudenberg and Maskin (1986) and Fudenberg, Levine, and Maskin (1994), among others.

[2] See Table 6 in the Appendix.

[3] See Connor (2007), as well as Marshall and Marx (2014). Harrington (2006) and Levenstein and Suslow (2006, 2011, 2014) summarize various stylized facts about real-world cartels.

and a rare opportunity to study the internal organization of present-day cartels. Our immediate research question is the following: Why did some cartels survive for a decade whereas others collapsed after only a few years? For example, the cartel for vitamin C collapsed in 1995, but those for vitamin A, vitamin E, and beta carotene continued operations until the prosecution in 1999 (see Figure 1). These four are among the largest of the 16 markets, and their data are the most comprehensive; hence, our empirical analysis focuses on them. In particular, vitamin C is the most interesting market because its cartel was the only one (among the four) that broke up without antitrust enforcement.

Figure 1: Cartels and Vitamin Prices



Note: Roche's monthly average U.S. transaction prices ($/kg) are rescaled with January 1995 as the base period for comparison across the four vitamin categories. The vitamin C cartel collapsed spontaneously after August 1995, whereas the other three cartels continued operations until 1999, when the U.S. government prosecuted them. See section 3 for details.
Source: Roche ROVIS data from Roche Data Books *cit. in* "Expert Report of B. Douglas Bernheim," *In Re: Vitamins Antitrust Litigation*, MDL No. 1285, Misc 99-0197. We reverse-engineered the price data by digitizing Figures 9-1 through 9-4 of this expert report.

Estimating a fully specified repeated game would appear to be a daunting task at first glance, but the actual steps are simple and straightforward. We analyze each market using a standard repeated-game framework in which the expected payoffs from complying with the cartel agreement must exceed those from non-compliance for the agreement to be self-enforcing (section 2).[4] The details of the model, including lagged perfect monitoring and trigger strategies with Nash reversion as punishment, are based on the direct evidence from court documents (section 3). We also explain the historical context of Chinese state-owned enterprises (SOEs), which became a major competitive fringe in the vitamin C market and contributed to the demise of the cartel. Bernheim's (2002a) report contains 3,140 pages with many graphs on prices, costs, production, and market shares, from which we reverse-engineered quantitative information (section 4). We use this data set to estimate a simple demand-and-supply model that underlies the stage-game profits, which in turn allow us to quantify the cartel's incentive compatibility (IC) constraint (section 5).

We find Roche's incentive to collude diminished significantly at the time of the vitamin C cartel's actual breakup in 1995. By contrast, nobody's IC constraint was binding in the three other markets (vitamin A, vitamin E, and beta carotene). Thus, our model explains the life and death of the four major vitamin cartels. This finding is robust to a range of alternative model specifications, including endogenous fringe supply, renegotiation of quotas, non-monopoly pricing by the cartel, adaptive expectations, functional forms of the demand curve, and long-term contracts (section 6).

After obtaining the measure of collusive incentives, we ask our second question: What determines the cartel stability? Many existing studies have asked this important question, and our contribution is in quantifying the effects of demand, fringe supply, and market structure on the IC constraint through two sets of fully dynamic counterfactual analyses. First, we closely investigate the breakup of the vitamin C cartel by re-computing the IC constraint under hypothetical scenarios in which (i) demand had grown more than it actually did in the mid 1990s and/or (ii) fringe exports from China had stopped growing in the mid 1990s. Results suggest the demand-side forces were at least as important as the fringe-supply factor for the collapse of the cartel (section 7.1).

The second counterfactual experiment (section 7.2) studies the effect of market structure on cartel stability by simulating a hypothetical merger between Takeda and BASF, the second and the fourth largest suppliers of vitamin C, in 1990 (i.e., before the beginning of their cartel) instead of 2001 when Takeda sold its vitamin business to BASF in reality. We use the U.K. Competition Commission's (2001) 190-page merger review to emulate the actual transaction. Theoretically, the effect of facing fewer competitors on collusive incentives is ambiguous because both cooperative gains and deviation gains can increase. Our results suggest the former effect dominates the latter in the current setting, considerably relaxing

[4] We say "non-compliance" rather than "deviation" because the vitamin C cartel collapsed in reality, whereas "deviations" are off-path by definition.

the IC constraint and making the vitamin C cartel sustainable at least until 1999, the year of the prosecution. The "coordinated effect" of this merger is large.[5] For example, the average 1998 price would have increased by merely 1.7% if we consider only the "unilateral effect" (i.e., an increased markup under oligopolistic competition with fewer firms), but the cartel could have raised it by 18.1% if we incorporate the coordinated effect as well. We decompose the effect of merger on the IC constraint into four theoretical components, and find the merger would have relaxed Roche's IC constraint primarily through more generous quota allocation among fewer firms. To our knowledge, this study is the first to quantify and decompose the coordinated effect based on a standard theoretical model.

Thus, we regard the theory-based measurement of the incentive to collude (and the effects of its determinants, including market structure) as our primary contribution. In the course of our analysis, we also show our simple model could explain the life and death of real-world cartels, thereby demonstrating the usefulness of one of the most fundamental predictions of repeated-game theory that cooperation is self-enforcing if and only if it is incentive compatible (i.e., the future is sufficiently important).

Literature Collusion is such a central topic in both repeated-game theory and IO that we refer the reader to Stigler (1964), Mailath and Samuelson (2006), Porter and Zona (2008), and other introductions and surveys. Our review focuses on papers that (i) consider repeated games explicitly and (ii) contain substantial empirical elements. According to these criteria, the most closely related papers are Porter (1983), Ellison (1994), Harrington and Skrzypacz (2007 and 2011), and Asker (2010).

Both Porter (1983) and Ellison (1994) studied an American railroad cartel from 1880 to 1886 to test theoretical predictions about the timing of price wars, focusing on the regime switching between collusive and competitive periods under demand uncertainty and imperfect public monitoring.[6] By contrast, we study a simpler environment with no (or limited) uncertainty and perfect monitoring.[7] Our primary goal is to measure the incentive compat-

[5]In antitrust, "coordinated effect" typically refers to the increased probability of tacit collusion, whereas economic theory does not necessarily distinguish between tacit and explicit collusion. We interpret our estimates as an upper bound of what colluders can achieve with or without communications.

[6]They test the theoretical models of Green and Porter (1984) and Rotemberg and Saloner (1986).

[7]Monitoring technology is divided into perfect and imperfect. *Perfect* monitoring means each player can observe each other's actions perfectly, whereas *imperfect* monitoring means they observe noisy signals. We can further classify imperfect monitoring into public and private. Monitoring is *public* when each player observes a common signal, such as the price of a commodity (e.g., crude oil and the OPEC). Monitoring is *private* when each player observes a private signal (not observed by other players), such as the price and quantity of differentiated products that are determined bilaterally in confidential meetings. Sometimes people use the words "perfect" and "public" almost interchangeably because perfect monitoring implies players observe the action profile, which necessarily constitutes a common signal.

ibility of the cartel scheme and how market structure affects it.

Harrington and Skrzypacz's (2007 and 2011) works are primarily theoretical, but they closely examine the case of the lysine cartel and construct a theoretical model to explain the actual strategies the colluding firms played.[8] We share their spirit in studying another high-profile antitrust case in recent decades; our empirical approach is more direct in the sense that we estimate a structural model of demand and supply in conjunction with a repeated game, by exploiting the abundance of quantitative and qualitative evidence.

Asker (2010) studies a bidding ring in auctions of collectible stamps by using rich data on its internal organization. Bid rigging is a different form of collusion and more closely connected to the auction literature, but we share the common theme of collusion as well as the empirical approach of using direct evidence from criminal cases.[9]

In this literature context, we regard the combination of an oligopolistic repeated game and a structural analysis of demand and supply as the defining characteristic of our work.

2 Model

We present our baseline model of a cartel in a repeated game framework. Our key assumptions are based on the facts and data from the court documents (sections 3 and 4). We defer the exposition of empirical specifications to section 5, and robustness checks to section 6.

2.1 Baseline Model

A fixed set of firms $I = \{1, ..., n\}$ play the following repeated game. We treat a specific category (such as vitamin C) as an independent market, and do not model "multi-market contact" in which a deviation in one market is punished in all of the other markets.[10] The cartels for vitamin A, vitamin E, and beta carotene outlasted the vitamin C cartel by four years. That is, a failure in one market did not trigger punishment in the others. We assume I is fixed and do not model entry or exit, because the profile of the major vitamin producers during the 1980s and 1990s is stable, and so is that of the participants of each cartel.

[8] De Roos (2006) estimates a standard model of demand and supply for the lysine case. Other recent case studies based on repeated games include Greif (2006), and Kandori and Obayashi (2014).

[9] Other recent structural empirical studies include Clark and Houde's (2013) work on a retail gasoline cartel, Miller and Weinberg's (2015) work on tacit collusion and merger in the beer industry, and Kawai and Nakabayashi's (2015) work on the detection of bid rigging in government procurement auctions.

[10] The term "multi-market contact" is used in various ways. For example, some literature use it to refer to any situations in which firms compete or collude in more than one markets. Here we use the term more specifically to refer to the possibility that a deviation in one market will cause punishment in another. See Matsushima (2001), Kobayashi and Ohta (2012), and Sekiguchi (2015) for the theoretical background.

In each period t, economic state X_t is realized, which is observable to each firm. Each firm decides its supply $q_{i,t}$ given its private history h_i^t (to be defined). Given the total supply $Q_t = \sum_{i \in I} q_{i,t}$, the market price $P_t = \frac{dP}{dQ} \times (Q_t + Q_{fri,t} - X_t)$ is realized, where $\frac{dP}{dQ}$ is the slope of the (inverse) demand curve,[11] X_t is the observable demand shifter, and $Q_{fri,t}$ is the supply from the competitive fringe suppliers. Firm i's profit is given by

$$\pi_{i,t} = (P_t - c_{i,t})\, q_{i,t}, \tag{1}$$

where $c_{i,t}$ is its constant marginal cost.[12]

Prices of the same product charged by different firms are closely correlated and practically identical throughout the sample period (see Figure 9 in the Appendix), which suggests the market price is common and contemporaneously observed. By contrast, players observe the quantity profile with L periods of lag: $\left((q_{j,\tau})_{j \in I}, Q_{fri,\tau} \right)_{\tau \leq t-L}$. The EC's (2003) criminal investigation revealed the cartel members exchanged internal sales records at quarterly meetings and used government statistics to verify and monitor each member's adherence to the output quota agreement, which suggests their monitoring is perfect with a three-month lag. We set the period length of our model to one month to match the frequency of the price data (i.e., $L = 3$). Under this physical and institutional setup, firm i's private history at the beginning of period t is

$$h_i^t = \left((q_{j,\tau}, Q_{fri,\tau})_{j \in I, \tau \leq t-3}, q_{i,t-2}, q_{i,t-1}, (P_\tau)_{\tau \leq t-1} \right),$$

and the public history is

$$h^t = \left((q_{j,\tau}, Q_{fri,\tau})_{j \in I, \tau \leq t-3}, (P_\tau)_{\tau \leq t-1} \right).$$

The firms know past X_t's and $c_{j,t}$'s, and rationally expect future X_t's and $c_{j,t}$'s. That is, $\left(X_\tau, (c_{j,\tau})_{j \in I} \right)_{\tau=1}^{\infty}$ is common knowledge, and hence we omit them from the history.[13] These assumptions are rooted in the fact (see section 3) that (i) the vitamin demand exhibits a steady growth trend and (ii) the production technology is mature and common across firms (up to some cost heterogeneity). As a robustness check, we also implement an alternative

[11] We specify a linear demand in section 5 and discuss an alternative, log-linear specification in section 6, but our identification of dP/dQ does not rely on functional-form assumptions.

[12] Bernheim (2002b) and other expert reports reveal the firms' internal cost data, which exhibit constant unit cost at any quantity or (nameplate) "capacity" utilization rates. See section 4 for further details.

[13] Our empirical analysis in section 5 allows dP/dQ to vary over time as well, and assume the firms form rational expectations over future dP/dQ_t as well.

model with adaptive expectations about demand (section 6). Regarding the post-sample period, we assume the demand and the costs remain constant at their end-of-sample values.

Let $q_{i,t}^N$ be the unique equilibrium quantity in the static Nash equilibrium given X_t and $c_{j,t}$. Because the evolution of X_t and $c_{j,t}$ does not depend on the firms' outputs, producing $q_{i,t}^N$ in each period t is an equilibrium (in the dynamic game).

Finally, EC (2003) reveals the cartel members held static expectations about $\{Q_{fri,t}\}_t$. That is, in each period t, they expect $Q_{fri,\tau} = Q_{fri,t-3}$ for each $\tau \geq t$. We revisit this point for further theoretical investigation in section 2.3, and provide factual details in section 3.

2.2 Equilibrium Concept

We consider the following equilibrium based on trigger strategies. The cartel members are supposed to supply their respective quota allocations, $\left(\bar{q}_{i,\tau|t}\right)_{i\in I,\tau\geq t}$, given the rational expectation of $\left\{X_\tau, \{c_{j,\tau}\}_{j\in I}\right\}_{\tau\geq t}$ and the static expectation $Q_{fri,\tau} = Q_{fri,t-3}$ for each $\tau \geq t$.[14] We write $\tau|t$ to indicate the expected future cartel production plan for period $\tau \geq t$ conditional on the static expectation $Q_{fri,\tau} = Q_{fri,t-3}$ formed as of period t. They agree on the quotas to maximize the total profit,

$$\sum_{i\in I}\left(\frac{dP}{dQ}\left(Q_\tau + Q_{fri,\tau|t} - X_\tau\right) - c_{i^*,\tau}\right) q_{i,\tau|t},$$

where i^* is a specific firm (e.g., Roche). Although $c_{i,\tau}$ differs across firms to some extent, the core technological process is common, and Roche is the undisputed leader of the industry as well as most of the cartels. Moreover, Roche's internal cost data are by far the most comprehensive and reliable of all cartel firms'. Hence, we compute *theoretical* cartel prices based on Roche's cost (i.e., as if each member shares the same cost structure), which turn out to match closely with the *actual* cartel prices in the data, in all of the four markets we analyze (section 5).[15]

We say non-compliance is confirmed in period τ if, given the government statistics, it becomes common knowledge that some firm did not produce $\left(\bar{q}_{i,s|t}\right)_{i\in I}$ in period $\tau - 3$ for the first time. That is, for each $s < \tau - 3$ and $i \in I$, we have $q_{i,s|t} = \bar{q}_{i,s|t}$, but there exists $i \in I$ with $q_{i,\tau-3|t} \neq \bar{q}_{i,\tau-3|t}$.

In each period t, given the expectation formed in period t, the firms agree to play the

[14] EC (2003) reveals the use of sales quota as a primary instrument for coordination, which appears to be a common collusive practice in many homogeneous-good markets according to Marshall and Marx (2014).

[15] As a robustness check, we also investigate other pricing rules in section 6.

following strategy for period $\tau \geq t$: (i) if no non-compliance is confirmed previously in period τ, then each firm sells $q_{i,\tau|t} = \bar{q}_{i,\tau|t}$; and (ii) if some non-compliance is confirmed in some period $s \leq \tau$, then each firm sells a static Nash equilibrium quantity $q_{i,\tau} = q^N_{i,\tau|t}$. Let us call this strategy the "trigger strategy."[16]

For notational convenience, let $\pi^C_{i,\tau|t}$ be firm i's period profit under the quota $\left(\bar{q}_{j,\tau|t}\right)_{j\in I}$, $\pi^D_{i,\tau|t}$ its optimal non-compliance profit against $\left(\bar{q}_{j,\tau|t}\right)_{j\neq i}$, and $\pi^N_{i,\tau|t}$ its static Nash equilibrium profit. Complying with the cartel agreement from period τ on gives firm i the payoff of

$$V^C_{i,\tau|t} = \sum_{s\geq\tau} \beta^{s-\tau}\pi^C_{i,s|t}, \tag{2}$$

where $\beta \in (0,1)$ is the discount factor. When firm i does not comply in period τ, the optimal deviation payoff is

$$V^D_{i,\tau|t} = \sum_{s=\tau}^{\tau+2} \beta^{s-\tau}\pi^D_{i,s|t} + \sum_{s\geq\tau+3} \beta^{s-\tau}\pi^N_{i,s|t}, \tag{3}$$

because no punishment is conducted until the government statistics verify i's non-compliance three months later (i.e., $L = 3$).

In each t, given $V^C_{i,\tau|t}$ and $V^D_{i,\tau|t}$ for $\tau \geq t$, if there exist $i \in I$ and $\tau \geq t$ for which $V^C_{i,\tau|t} < V^D_{i,\tau|t}$, then it becomes common knowledge among the players as of t that some player will deviate in period τ and the prevailing actions will be a static Nash equilibrium from period $\tau + 3$. The situation becomes the same as in a finitely repeated game, and hence each firm deviates in period t. Consequently, the trigger strategy is an equilibrium (given the firms' expectations in period t) if and only if

$$\min_{i\in I,\tau\geq t} \left(V^C_{i,\tau|t} - V^D_{i,\tau|t}\right) \geq 0. \tag{4}$$

Recall that we assume players form static beliefs about $Q_{fri,\tau}$. In the vitamin C market, which is the only case (among the four cases we scrutinize) in which the cartel collapsed before the prosecution, $Q_{fri,t}$ increased over time in a staggered manner, and hence players revised the expectation of $Q_{fri,\tau}$ increasingly more pessimistically over time. Eventually, the cartel must break up in the period in which (4) is violated for the first time. This setup implies that

[16] Given the static expectation of fringe supply, firms in period t believe their expectation of period-τ quantity, $q_{i,\tau|t}$, is going to equal the actual quantity, $q_{i,\tau|\tau}$. If their belief is correct, this description is exactly the trigger strategy: given agreed production $(\bar{q}_{i,\tau})_{\tau\geq t}$, a deviation will lead to the repetition of the static Nash equilibrium.

for the market in which the cartel collapsed, we do not interpret the entire sequence of the play as part of a *single* equilibrium strategy as in Porter (1983) or Ellison (1994). Instead, we focus on the actual history of bad news (e.g., an unprecedented growth of the vitamin C exports from China) and propose the following interpretation. At the beginning of the cartel, players expected the future environment would make the trigger strategy an equilibrium (in particular, that no breakup would happen on the equilibrium path). At some point in the subsequent periods, however, the previously unforeseen negative news about $Q_{fri,t}$ arrives and forces them to realize the cartel agreement is no longer an equilibrium. The members switch to the repetition of static Nash equilibrium as a consequence.[17]

2.3 Assumptions behind the Theoretical Model

In the preceding model, the following three assumptions about strategy are particularly important. We discuss their theoretical background in this subsection. We defer the exposition of their factual foundations to section 3.

Static Expectations about Fringe Supply

The first key assumption is the cartel's static expectation about $Q_{fri,t}$. The expert reports and other records suggest the vitamin C cartel broke up partly because of the sudden expansion of the Chinese SOEs, which doubled and tripled the export quantity within a few years in the middle of the 1990s, as a result of a series of geopolitical events. The growth of fringe supply reduced the residual demand left for the cartel as well as its members' continuation payoffs, eventually violating their incentive compatibility constraint in (4). Thus, exactly how much of this fringe expansion the cartel foresaw affects the timing of its collapse.

Our motivation is two-fold. The immediate reason we use static expectations is factual: EC (2003) suggests the cartel agreed on quota allocations in each year based on the optimistic forecast of fringe supply at the previous year's level (see section 3). The other reason is theoretical. Suppose each player belongs to one of the two possible types, sophisticated and naive. The sophisticated type rationally expects that, once the Chinese SOEs start growing, they will eventually violate (4). Meanwhile, the naive type holds the static belief and has to update it with news. In the best equilibrium for the cartel, sophisticated members will optimally choose to pretend to be naive with a high probability for a long time, as long as some higher-order uncertainty exists about the types of other members. Moreover, this result holds even if the ex-ante probability that all firms are sophisticated is high.

[17] We study the possibility of renegotiation in one of the robustness checks.

The intuition is the same as in Kreps, Milgrom, Roberts, and Wilson (1982).[18] The game is finitely repeated in their model; the game is effectively finitely repeated in our model as well, because the fringe supply will eventually violate (4). Likewise, the naive type in their context means the commitment type to cooperation (in the finitely repeated game); the naive type in our model means the "commitment" type that follows the cartel agreement until (4) is violated in its static belief. In both cases, the pretense of naiveness can extend the length of cooperation (cartel) periods, whereas the cooperation (cartel) has to end immediately if it becomes common knowledge that every player is actually sophisticated.

Hence, the following alternative specification yields an equilibrium behavior similar to our baseline model. Every player assigns a low ex-ante probability to the future contingency in which China invents the low-cost production method and its SOEs dominate the world market, until they actually do. Once the fringe supply starts growing, the cartel firms play the game of incomplete information (i.e., sophisticated vs. naive) in the above.

Perfect Monitoring with Delay

The second important assumption is that no punishment is conducted until non-compliance is verified with the government statistics. We assume X_t and P_t are commonly observed at the beginning and the end of period t, respectively. Because $P_t = \frac{dP}{dQ}\left(Q_t + Q_{fri,t} - X_t\right)$, the reader may notice the theoretical possibility that each firm should be able to infer Q_t from P_t, $\frac{dP}{dQ}$, $Q_{fri,t}$, and X_t by the end of period t, thereby detecting its rival's non-compliance without having to wait for the publication of the government statistics. In reality, however, the cartel members discussed the possibility that someone had over- or under-achieved the quota (and the need for minor adjustments at the end of calendar year) only at the quarterly meetings, with the third-party verification at hand. That is, they did not act immediately on such privately inferred actions of each other.

How should we reconcile these two features of the model and the reality? Theoretically, we can reformulate the inverse demand as $P_t = \frac{dP}{dQ}\left(Q_t + Q_{fri,t} - X_t\right) + \varepsilon_t$, where ε_t is an i.i.d. noise with $\mathbb{E}\left[\varepsilon_t\right] = 0$. Given the full support of ε_t, individual firms can no longer identify Q_t from P_t, and they can rationally expect that the other firms have followed the equilibrium strategy after any P_t. Because each firm's payoff is linear in P_t (given $q_{i,t}$, which is known to firm i), extending our model to incorporate this mean-zero error, ε_t, does not change our analysis. Thus, we omit ε_t from the baseline theoretical model for simplicity.[19]

[18] See also Kreps and Wilson (1982) and Milgrom and Roberts (1982).

[19] A theoretical work by Fudenberg, Ishii, and Kominers (2014) suggests more sophisticated methods of punishment when evidence arrives with delay, but we keep our model closely tied to the actual internal

Punishment by Nash Reversion

The third assumption we wish to elaborate on is that the punishment is an infinite repetition of the static Nash equilibrium. We have chosen this specification because the cartel members explicitly stated their understanding of the agreement as such.

Theorists have proposed severer forms of punishment, thereby suggesting the possibility that the equilibrium payoffs under the trigger strategy such as ours might not represent the theoretically best-performing cartels. For example, Abreu (1986) constructs "stick-and-carrot" equilibrium. Fudenberg and Maskin (1986), Abreu, Pearce, and Stacchetti (1990), and Fudenberg, Levine, and Maskin (1994) show players can implement a severe punishment incentivized by a continuation payoff. However, the data on vitamin C prices after the cartel's collapse indicate the market was stable, and do not exhibit patterns indicating the movement of continuation payoffs as these advanced theories predict.[20]

In section 6, we alter and extend our baseline model in several directions for robustness checks. Before engaging in further theoretical discussions, however, let us first establish the facts from the court documents and other antitrust sources from the United States, the United Kingdom, and the European Commission.

3 The Vitamin Cartels, 1990–1999

Vitamins represent an ideal natural laboratory for an empirical analysis of cartels. This section explains the industry background, the cartels, and China.

3.1 Industry Background

Each vitamin constitutes a separate product market because of its unique physiological properties and unique production processes. For example, humans cannot substitute vitamin A for vitamin C, and manufacturing facilities of vitamin E cannot produce beta carotene.[21]

Within each category, bulk vitamins are homogeneous goods. Buyers (e.g., farmers, blenders, and food and beverage manufacturers) choose vitamin suppliers based on prices, not brands.[22]

organization of the vitamin cartels as reported in EC (2003).

[20] This observation does not preclude the possibility that the distribution of continuation payoffs depends on strategies. Nevertheless, the post-cartel price data exhibit stationary patterns for at least a few years.

[21] Table 7 in the Appendix lists raw materials and intermediate inputs for each vitamin.

[22] Figure 9 in the Appendix confirms this observation by showing firm-level prices of four specific vitamin

Geographically, markets are global because transport costs are low relative to value, and independent traders engage in cross-border arbitrage. The fact that the price fixing required international cartels further confirms the global scope of vitamin markets.

Demand

Vitamins are defined as organic substances that (i) are natural components of foods, (ii) are essential in small amounts for normal physiological function, (iii) are not synthesized by the host (i.e., human or animal) in adequate amounts, and (iv) cause a specific deficiency syndrome when absent.[23] Hence, avoidance of deficiency symptoms is the primary nutritional purpose of vitamin intakes. Broader health benefits for humans have not been scientifically proven, but "educational marketing" has promoted the perception of its benefits among consumers.[24] By contrast, animal nutritionists conduct sophisticated cost-benefit analyses to determine the optimal mix of vitamins for each species. Thus, each vitamin has a different demand base. More than 90% of vitamin C and beta carotene are for human use, whereas 87% of vitamin A and 73% of vitamin E are for animals.

All vitamin markets experienced a steady growth of demand during the 1980s and the 1990s. Industry experts regard the following factors as key determinants of demand growth: population of humans and animals, GDP per capita (countries above a certain income level tend to exhibit rapid expansion of markets for health products), the general public's perception of health benefits, and sophistication of animal husbandry.

Buyers of bulk vitamins include farmers, their cooperatives, local blenders who produce pre-mixed vitamin cocktails, manufacturers of foods and beverages, drugs and cosmetics firms, and other firms that use vitamins for miscellaneous technical purposes. The class actions involved more than 4,000 plaintiffs and more than 9,000 purchasers on record. The EC report recognizes a few "large" buyers, primarily in the vitamin C market, who used it as an antioxidant additive for foods and beverages. But even the largest buyer, Coca-Cola, accounted for only 2.14% of the aggregate sales during the cartel period. Thus, buyer concentration is generally low, and price cuts cannot be kept secret.[25]

---

products (Vitamin E 50% Adsorbate Feed Grade, Vitamin E Acetate Oil USP, Ascorbic Acid 100% USP, and Beta Carotene 30% Fluid Soluble), which are highly correlated and close to each other in levels.

[23] Choline and carotenoids are not among the 13 substances generally recognized as vitamins, but are considered below because they were an integral part of the antitrust litigation.

[24] According to Bernheim (2002, Appendix E, pp. 9–10), "vitamins for human nutrition lack a fully accepted fact-based correlation between consumption and objective performance measures" and "except when used as chemical antioxidants, vitamins are included in foods for their label value."

[25] See Connor (2007, pp. 254–256) for details.

Supply

Practically all major suppliers of vitamins (21 in total) joined the cartels, with each vitamin category typically dominated by three or four firms. Table 1 compiles the market shares of selected companies in the cartels. Hoffmann-La Roche (henceforth Roche) is a Swiss drug maker that pioneered the mass production of vitamins in the mid 20th century and the undisputed industry leader. In the 1970s, Badische Anilin und Soda Fabrik (BASF), a German chemical giant, successfully imitated Roche's process in several categories. Rhône-Poulenc (RP), a French chemical maker, also began producing a few vitamins. Together, these European "big three" formed the core of the global cartels in the 1990s.

Table 1: Global Market Shares (%) by Category in Early 1990s

Market: Firm	A	B1	B2	B3	B4	B5	B6	B9	B12	C	D3	E	H	Caro-tinoids	All
Roche	48	44	54	–	–	36	49	39	–	46	43	46	45	83	46
BASF	30	2	30	–	15	21	3	–	–	7	13	28	–	16	17
RP	21	–	–	–	–	–	–	–	62	–	–	13	–	–	8
Takeda	–	31	3	–	–	–	12	23	–	26	–	–	–	–	7
Eisai	–	–	–	–	–	–	–	–	–	–	–	12	–	–	2
Daiichi	–	–	–	–	–	29	12	–	–	–	–	–	–	–	1
E. Merck	–	–	–	–	–	–	5	–	–	10	–	–	10	–	2
Hoechst	–	–	–	–	–	–	–	–	7	–	–	–	–	–	1
Other members	–	–	–	86	75	–	–	35	–	–	44	–	42	–	9
Cartel total	99	77	87	86	90	86	81	97	69	89	100	99	97	100	93
Non-members	1	23	13	14	10	14	19	3	31	11	0	1	3	0	7

Source: Connor (2007, 2008).

Takeda Chemical Industries (Takeda) is a bulk-chemicals arm of the largest pharmaceutical firm in Japan. It began volume production and exports of vitamins in the 1950s. By the 1990s, Takeda had become the second or third largest global manufacturer of several vitamins, followed by Eisai, Daiichi Pharmaceutical, and other domestic rivals. By contrast, American companies had ceased to be major suppliers by the 1980s, despite their early involvement in vitamins.[26] Their facilities were either closed or sold to European makers.

Since 1980, the production technologies of vitamins have not undergone any major changes. The only exception is the invention and commercialization of a new method to produce vitamin C in China, which we will describe in section 3.3. No major entry or exit occurred during the cartel period except for fringe firms.

[26] For example, Pfizer and American Home Products manufactured vitamins. The first research director of Merck focused on vitamins in 1933, which contributed a large share of the company's total sales in the late 1930s. This U.S.-based Merck is not to be confused with E. Merck, its German parent before World War I. The latter was a cartel member in the 1990s.

3.2 The Cartels

Twenty-one manufacturers joined one or more of the 16 conspiratorial groups, each of which regularly met to set quotas and target prices, exchange internal sales records, and monitor implementation of the quotas. Roche, the industry leader, organized 14 of the 16 cartels, with BASF as its main partner in crime (12 cartels), followed by Takeda (5 cartels) and RP (3 cartels).

Beginning

On June 7, 1989, the heads of the vitamins divisions of Roche and BASF met in Basel, Switzerland, to start cartels for vitamins A and E. RP's head of the Animal Nutrition division joined two months later at another meeting in Zurich. They agreed to freeze market shares at the 1988 levels for the foreseeable future and to split their predicted 1990 sales proportionally to these quotas. This practice came to be known as the "budget meeting" in every late summer or fall. These top-level annual meetings were supplemented by middle-level meetings held at the quarterly frequency among the heads of worldwide product marketing, who also exchanged progress reports on sales volumes. At the lowest level, regional product marketing managers also met four times per year to monitor regional quotas, assess trends in demand and supply, and make small adjustments to prices in local currencies. With minor variations, these basic rules and organizational structure for vitamins A and E applied to all of the other cartels.

In 1990, the three original members recruited Hoechst, a German maker of vitamin B12, and Eisai of Japan, which was the only major producer of vitamin E besides Roche and BASF, to solidify the two existing schemes for vitamins A and E, and to form four more cartels for vitamin B12, two carotenoids (beta carotene and canthaxanthin), and premixes (i.e., customized cocktails for animal feed, consisting mainly of vitamins A and E). These six cartels were operational in 1990. To ensure cooperation of new members, the leading members would often cede some of their historical market shares. Furthermore, Roche contacted Daiichi, E. Merck, and Takeda to cartelize the markets for vitamins B1, B2, B5, B6, B9, C, and H. Takeda also agreed to become a liaison with smaller Japanese makers, including Sumitomo Chemical, Tanabe Pharmaceutical, and Kongo Chemical. By early 1991, all 14 of the Roche-led cartels were successfully raising prices of bulk vitamins.

Operations

Monitoring was close to perfect. Third-party data supplemented the monthly and quarterly exchange of internal sales records. The members knew the location of each member's plants, and typically a country would have only one producer for a given product. Thus, international trade statistics from the respective governments could serve to verify self-reported sales information. This reasonably transparent monitoring environment facilitated the implementation of the quota agreement.

The cartels did not specify punishment, but the implicit threat was that the members would revert to competitive pricing and the collusion would break down indefinitely, which is typical of cartel agreements in the real world (c.f., Harrington 2006, Marshall and Marx 2014). Such a shared understanding was occasionally aired, for example, by E. Merck in the vitamin H market and by Roche, E. Merck, and BASF in the vitamin C market. EC (2003) records that "the three European producers presented Takeda with an ultimatum: unless it agreed to cut back its vitamin C sales, they would withdraw from the agreement" (p. 44, emphasis added). Throughout the 1990s, nothing like a full-blown price war was observed, except when several of the cartels collapsed permanently.

Most cartels set different target prices for three regions (Europe and the Middle East, North America, and the rest of the world), but the geographic price spreads were designed to be less than 10%. The 10% rule reduced arbitrage opportunities for third-party traders.

End

The 16 cartels ended in one of two ways. Six of them fell apart in 1994 or 1995, apparently because of the entry and expansion by fringe producers and/or the difficulty of aligning the interests of relatively many participants. Chinese producers, mostly state-owned enterprises (SOEs), significantly increased markets shares in vitamins B1, B6, B9, and C (see below for details). Il Sung, a Korean maker, was the fringe producer of vitamin H. Finally, Archer Daniels Midland (ADM), an American food-processing and commodities-trading company, entered the vitamin B2 market by acquiring facilities from Coors Biotech, a subsidiary of an American beer company.

By contrast, the 10 other cartels were operating smoothly until private and government investigations in the United States forcibly terminated them in 1998 or 1999. American investigators first learned about the vitamin cartels in late 1996 from ADM, which was cooperating with the DOJ in its investigation of a separate case concerning the citric acid cartel. In March 1997, the FBI interviewed Dr. Kuno Sommer, then president of Roche's

Vitamin and Fine Chemicals division, who denied the existence of any cartel. In March 1998, Boies & Schiller, a law firm, filed a civil price-fixing suit in the U.S. District Court for Dallas, Texas, on behalf of several direct purchasers of bulk vitamins. These allegations were forwarded to the DOJ and the FBI. In the summer of 1998, Lonza, a Swiss producer of vitamin B3, and Bio-Products, an American producer of vitamin B4, began cooperating with the FBI.

Finally, in January 1999, RP, the world's third-largest vitamin firm, applied for (and was granted) amnesty under the DOJ's relatively untested Corporate Leniency Program. RP's managers were rumored to be required to attend the last cartel meetings in February 1999 and tape-record it. Within two months, both Roche and BASF pled guilty and agreed to pay record-breaking fines of $725 million in total. RP saved more than $100 million in U.S. fines, but its main motivation for defection was that its drugs division (which was larger and more influential than the bulk chemicals division) wanted to merge with Hoechst, a German rival. RP's top management discovered the vitamin conspiracies in this process and decided to inform the authority in pursuit of regulatory approvals for its planned merger. The American and the European agencies did approve this merger, which created Aventis, one of the two predecessors of Sanofi-Aventis (currently Sanofi).[27] These historical developments raise an interesting question regarding how antitrust agencies can jointly optimize the enforcement of cartel and merger policies.

3.3 Competitive Fringe in China

Chinese producers played a major role as the competitive fringe in the breakup of the vitamin C cartel. This subsection provides their institutional background. Our main finding is that the unusual success of Chinese firms in the vitamin C market is a result of a series of unlikely events, including (i) the discovery of a new and efficient production method in the mid 1980s, (ii) a military event that disrupted the supply in 1992, and (iii) the government's promotion of entry and investment along with economic liberalization in the mid 1990s.

China has not always been a socialist economy; it became one in the 1950s. After the death of Mao Zedong in 1976, however, Deng Xiaoping led economic liberalization. Table 2 lists key events, of which the most relevant are the dissolution of the Soviet Union in 1991, Deng's Southern Tour Speech in 1992, which officially endorsed private enterprises,

[27] The underlying industry trend was that most drug makers were shedding non-drug operations at the turn of the century, in an effort to increase profitability (Yamaguchi 2000). In the same vein, Takeda sold its vitamin business to BASF in 2001 (see section 7.2).

Table 2: Historical Background of SOEs and Economic Liberalization in China

Year	Event
I. Central Planning	
1949	People's Republic started
1950–53	Korean War
1953	Soviet-style central planning began (First Five-Year plan):
	All private enterprises reorganized into the public sector by the end of the 1950s
1976	Mao Zedong died
II. Economic Liberalization	
1978	Deng Xiaoping starts economic liberalization policy
1980	Decentralization of budgetary control to local governments
1986	Legalization of private enterprises (with 8+ employees)
1991*	Soviet Union dissolved
1992*	Deng's Southern Tour Speech endorsed private enterprises
1994*	Liberalization of prices and commerce complete
	Modern corporate law and labor law became effective
1995*	SOEs started making losses collectively;
	Restructuring and layoffs at SOEs permitted
1996	Last Five-Year Plan with production quantity targets
III. Restructuring and Privatization of SOEs	
1997	Private sector endorsed as an "important part of socialist market economy"
	Privatization of SOEs endorsed under the slogan "Grasp the large; let go of the small"
2002	Privatization of small and medium local SOEs (and layoffs at large SOEs) mostly complete
	Accession to World Trade Organization

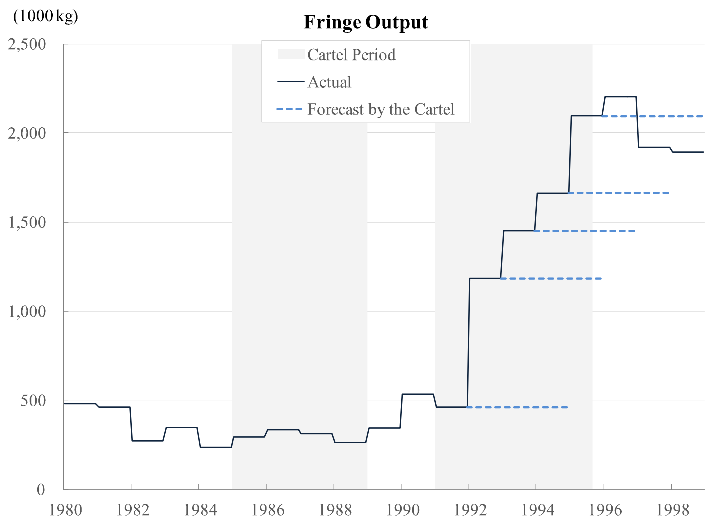
Note: * indicates key events that are related to the fringe expansion. See Chen, Igami, Sawada, and Xiao (2017).

the liberalization of prices and commerce in 1994, and the start of restructuring of SOEs in 1995.

Meanwhile, China's first patent law came into effect in 1985. Researchers at public-sector laboratories discovered and patented a new method involving two-step fermentation to manufacture vitamin C at a lower cost and smaller scale relative to the existing technology (the Reichstein one-step synthesis method). Despite this major invention, "Chinese vitamin C production had been of little importance internationally" until the 1990s (UKCC 2001). In the mid 1990s, however, an abrupt policy change occurred and "the Chinese central government had encouraged the improvement of the two-step process and its commercialization at various production sites," according to BASF's report to the UKCC.

Figure 2 shows the sudden growth of fringe supply in 1992. The event that triggered the government's promotion of vitamin C appears to be the Bosnian war in 1992, during which one of the NATO air strikes accidentally hit a large vitamin C factory that had supplied the former socialist bloc (Li 2002). In response to supply shortage, the Chinese government encouraged SOEs to increase production and supported the entry of private firms. As a result, more than 20 firms had entered the vitamin C market by 1995 (the exact number varies between 18 and 28, even within the same report by the UKCC), but most of them had disappeared by 2001 in what seems to be a typical industry dynamic of shakeout and consolidation. We profile these Chinese SOEs in the Appendix.

Figure 2: Sudden Growth of Fringe Supply



None of these events seemed to be precisely forecasted before 1992. For example, EC (2003) reveals that, even after three years of China's rapid growth in vitamin C exports, the cartel was allocating its 1995 quotas based on the premise that fringe supply would stop growing from its 1994 level. The dashed lines in Figure 2 correspond to these static expectations at various points in time.

Entry into the export market took time for regulatory reasons. The U.S. FDA had to inspect and approve any manufacturing plant in the world before its product could be used in pharmaceutical applications in the United States. Exports to the European Union require a Certificate of Suitability by the European Directorate for the Quality of Medicines. Given these institutional contexts, we characterize the Chinese producers in the 1990s as fringe suppliers with an exogenously determined time path of staggered entry and expansion, in our baseline model (sections 2 and 5), and as a lagged supply function (i.e., endogenous fringe) in some of the robustness checks in sections 6 and 7.

4 Data

4.1 Sources

The primary evidence for this study is a large amount of court documents and government reports that originated from the criminal investigations and civil litigation in America and Europe. As part of prosecution in 1999, the FBI and the DOJ obtained data on production and sales, personal records of cartel meetings, and individual depositions.

These antitrust enforcements were followed by multi-district litigation by class-action plaintiffs, which were subsequently consolidated at the U.S. District Court for the District of Columbia. On May 24, 2002, Dr. B. Douglas Bernheim, a Stanford economics professor, submitted a 339-page expert report with 2,801 pages of appendices on behalf of the lead plaintiffs. In response, defendants collectively served 16 expert reports on June 28, 2002, to which the plaintiffs further responded with a Reply Expert Report by William Landes, Hal Sider, and Gustavo Bamberger, as well as a Rebuttal Report by Bernheim, on July 17, 2002. These expert reports were included in jury trials in 2003 and therefore made publicly available, in principle.[28]

Another important source is the European Commission's antitrust enforcement decision to impose fines on the vitamin cartel members on November 21, 2001. The EC published an 89-page report in 2003, documenting the details of the internal organization of the cartels.

In addition, the U.K. Competition Commission published a 190-page report in July 2001 on the proposed acquisition by BASF of Takeda's vitamin businesses. This report does not directly investigate the cartels but conducts a thorough merger review and contains rich information on the industry background. This BASF-Takeda merger in 2001 motivates one of our counterfactual simulations, in section 7.

Finally, many economists participated in these litigation on both sides. Some of them went on to publish academic articles, including Connor (2007, 2008), Marshall, Marx, and Raiff (2008), and Marshall and Marx (2014). These studies constitute our secondary source of information and help us correctly interpret the primary evidence.

4.2 Variables

Bernheim (2002a) states that "the data available for damage estimation in this case are abundant and of unusually high quality" (p. 165). His analysis relies primarily on the

[28] In practice, a researcher has to visit the D.C. district court and manually investigate several thousand paper documents, because most of the court documents before 2005 have not been digitized.

internal data from Roche, the cartel leader. Roche's data set (ROVIS) contained "monthly weighted average prices (aggregated over customers) for a collection of vitamin products." He also had access to customer-specific transaction-level data from various plaintiffs and defendants, and hence both sides have cross-validated all information.

Prices

Bulk vitamins are homogeneous goods within each category, but multiple grades of concentration exist. For example, vitamin C was available in many different grades during the sample period,[29] but all of them are derivatives from the pure crystal form. Bernheim converts them to "100% basis" and aggregates the product-level sales to a single price index. EC (2003) reveals the cartel members coordinated prices according to the same "100% basis" calculations. Hence, this level of aggregation retains relevant information on pricing.

Costs

Bernheim (2002a) computes variable costs by using the annual worldwide "contribution margins" (i.e., profit margins) and the 100% basis average prices for each vitamin family reported in the Roche Data Books. He also received data on vitamin-level and product-level costs from several firms, but "the pattern of costs for other manufacturers is similar to the Roche Data Book, which I adopted whenever available" (p. 200), because of the quality and representativeness of Roche's data, which the defendants did not dispute. Some of their experts used Roche's data as well.

Despite their "unusually high quality," the average variable costs in these data have limitations in that they are accounting measures based on engineering estimates, which could potentially differ from economic marginal costs. In the current empirical context, however, we have three reasons to accept Bernheim's unit-cost data. First, Bernheim was a trained economist on the plaintiffs' side, and hence had no incentive to inflate cost data (e.g., by manipulating depreciation accounting). Roche's unit-cost measure consists of labor, raw materials, and other intermediate inputs such as electricity (i.e., variable costs). Bernheim declares, "I did not include any fixed cost measures in my analysis" (p. 180).[30] Second,

[29] Bernheim's (2002a, p. 39) Table 6-3 lists the following 11 products: (1) Ascorbic Acid 100%, (2) Ascorbic Acid Compressible 90% USP, (3) Sodium Ascorbate USP, (4) Ascorbic Acid Compressible 95% USP, (5) Ascorbic Acid Compressible 97.5% USP, (6) Ascorbic Acid Coated 97.5%, (7) Niacinamide Ascorbate USP, (8) Calcium Ascorbate USP, (9) Ascorbic Acid Stabilized Feed Grade, (10) Ascorbyl Palmitate FCC, and (11) Other C Products. Note that all of the first 10 products are for human use except (9), which accounts for 0.7% of total U.S. sales in dollar value.

[30] Depreciation expenses typically accrue as part of SGA (selling and general administrative) expenses

Roche used this measure for its own production decisions. Bernheim notes, "since Roche apparently used this information when making contemporaneous business decisions, it is appropriate for my current purpose to take the data at face value" (p. 123). Third, data from the expert reports suggest physical production capacities were far from binding during the sample period (see below). Hence, the shadow price of capacity was unlikely to become a major concern.

Production

Bernheim (2002a) contains information on each firm's annual market shares. Similarly to the price indexes, all outputs are converted to 100% basis and aggregated at the vitamin level, which is the level at which the cartels defined and allocated quotas based on historical (pre-cartel) market shares.

Capacities

Bernheim (2002a) and other expert reports show "capacity utilization" data on several firms at the vitamin level, but these data did not play a major role in their analyses. We do not use them either. To clarify the conceptual and measurement issues surrounding capacity, distinguishing between three different notions of "capacity" (two conceptual and one empirical) is useful.

Conceptually, we are interested in either long-run or short-run capacities. By "long-run" capacity, we mean hard limits on output quantities defined by the physical sizes of production facilities. This first type of capacity never seemed binding during the sample period. Manufacturers conventionally planned for and maintained the sizes of facilities that were sufficient to accommodate a decade-long demand growth, which tended to be steady and predictable in this mature industry.

The second, "short-run" capacity corresponds to Kreps and Scheinkman's (1983) notion of quantity pre-commitment. In our empirical context with monthly spot-market prices, *monthly production schedules* (i.e., the smooth operations of chemicals plants require orderly work shifts and timely procurement of raw materials) would be the closest empirical counterpart to this concept.

Unfortunately, the reported "nameplate" capacities in the expert reports correspond to neither of these economic concepts, and would vaguely represent a noisy measure of medium-

rather than COGS (cost of goods sold) in a generally accepted accounting practice. The measurement and management of product-level "contribution margins" concern COGS, not SGA expenses.

run production plans (see Appendix A.5). Because different firms measure nameplate capacity based on different assumptions about the number of shifts per day or days per week the plant operates, Thomas McClymont, Roche's head of Production Coordination and Technology, stated in his deposition that "capacity is an opinion" whereas production is a fact.[31] Thus, we follow Bernheim in not using this ambiguous measure, and interpret the actual production as an embodiment of the firms' effective short-run production plans.

5 Empirical Analysis

Having studied the conceptual framework (section 2), the institutional background (section 3), and the data (section 4), we may now investigate the data patterns and provide economic interpretations. Our exposition in the main text focuses on the vitamin C market, but the Appendix contains the same analysis of the other markets (vitamin A, vitamin E, and beta carotene).

5.1 Data Patterns and Key Observations

Figure 3 shows quantities and prices in the vitamin C market. The left panel shows the worldwide production increased throughout the sample period. The four cartel members (Roche, Takeda, E. Merck, and BASF) restricted outputs in the mid 1990s, but the competitive fringe (mainly from China) started expanding in 1992. The right panel shows the price soared by more than 30% during the cartel period (January 1991 through August 1995) despite a virtually flat trajectory of unit cost.[32]

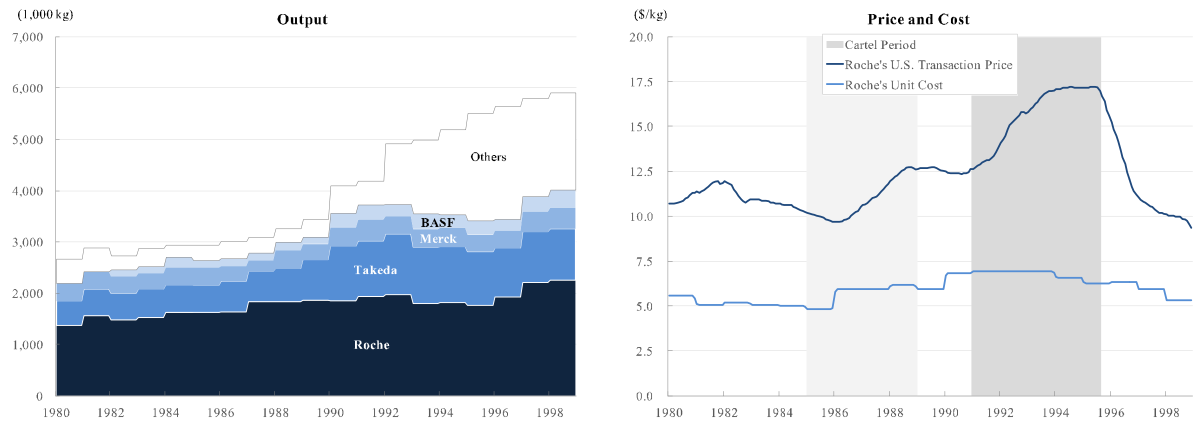
Based on these data patterns, Bernheim computed damages by using a reduced-form regression (lasso) of the price on a variety of demand and cost shifters. We need to take a step further by constructing a model of demand and supply because our objective is to explain the cartel's collapse by quantifying the firms' incentive to collude.

Two additional data patterns deserve further attention and foreshadow our identification strategies. The first observation concerns the demand side. The fact that both the price and the aggregate output increased in the mid 1990s suggests a steady growth of demand, which is consistent with the views of industry experts (e.g., Dr. Robert Speights, who wrote Appendix E of the Bernheim report). The second observation concerns the supply side. The

[31] Transcript of McClymont's deposition (p. 94), *cit. in* Landes, Sider, and Bamberger (2002, p. 49).

[32] Bernheim (2002a) suggests the possibility of another attempt to cartelize the vitamin C market between 1985 and 1988. We do not study this earlier period, because evidence is inconclusive and not publicly available. Nevertheless, our subsequent data analysis accounts for this possibility.

Figure 3: Price, Cost, and Quantity of Vitamin C



fact that the markup was sizeable even outside the cartel period (e.g., early 1980s) suggests the existence of market power. As we reviewed in section 3, every expert report characterizes bulk vitamins as homogeneous goods in a global market in which multiple large suppliers compete. Thus, the data reject Bertrand competition as a characterization of this industry, whereas the Cournot model could reconcile these pieces of evidence. See section 4.2 for our preferred interpretation of outputs and capacities in the spirit of of Kreps and Scheinkman (1983).

5.2 Demand and Supply

A linear demand model is a standard specification for simple homogeneous goods such as bulk vitamins (see section 6.5 for a log-linear version),

$$Q_t^D = \alpha_0 + \alpha_1 P_t + \alpha_2 X_t + \varepsilon_t, \tag{5}$$

where P_t is the price, X_t is a collection of demand shifters, and ε_t is a non-systematic component of demand shifters. A standard procedure is to regress Q_t on P_t and X_t to estimate the parameters $(\alpha_0, \alpha_1, \alpha_2)$. At least two instrumental variables (IVs) for P_t are available from Bernheim (2002a) and EC (2003): unit production cost, c_t, and a dummy variable, I_t, indicating whether the cartel was in operation.

However, this usual approach is confronted by two obstacles. One is a measurement problem of X_t. Bernheim (2002a) used human population, GDP per capita, and the number of slaughtered animals. De Roos (2001) used the number of news articles on the health benefits of vitamins. Despite such efforts to collect demand shifters, a precise and convincing measurement of X_t remains an elusive quest particularly for vitamin C and beta carotene, over 90% of which is used by humans, for which the scientific evidence of health benefits is lacking (see section 3.1).

The other problem is severe multicollinearity. Both Q_t and P_t increased in the mid 1990s, which implies an upward trend in X_t as well. This part of the sample exhibits the largest variation in data, to the extent that the estimates of αs become highly sensitive to small changes in the specification and functional forms of regression equations, frequently generating upward-sloping demand curves and other curiosities.

A more attractive alternative to such elusive quests is to exploit information and modeling assumptions on the supply side. From the criminal investigations and court documents we know exactly when the regime (i.e., I_t) switched between competition and collusion; hence, we know the appropriate period to apply the firms' first-order conditions (FOCs) for non-cooperative profit maximization. Section 5.1 showed the data patterns at odds with Bertrand competition and in favor of Cournot competition,[33] the FOC of which is

$$P_t + \frac{dP}{dQ} q_{i,t} = c_{i,t}, \tag{6}$$

where $\frac{dP}{dQ}$ is the slope of the inverse demand curve, $q_{i,t}$ is firm i's output, and $c_{i,t}$ is firm i's marginal cost. Our data contain measures of P_t, $q_{i,t}$, and $c_{i,t}$, which identify $\frac{dP}{dQ}$ and allow it to vary over time. Once we obtain the estimates of $\frac{dP}{dQ}$, the price coefficient in (5) is its inverse (i.e., $\frac{dP}{dQ} = \frac{1}{\alpha_1}$). Hence, we may calculate $\alpha_1 P_t$ and subtract it from Q_t to obtain the estimates of the *effective* demand shifter,

$$\tilde{X}_t \equiv \alpha_0 + \alpha_2 X_t + \varepsilon_t, \tag{7}$$

which is the sum of the non-price terms in (5), without relying on imperfect measures of X_t. The only identifying assumption is the additive separability of the price term.[34]

Every identification strategy has its benefits and costs. The benefit of this supply-side approach is two-fold: (i) no need to rely on functional-form assumptions on demand other than linear separability and (ii) no need to perfectly observe all demand shifters. The cost of this approach is also two-fold: (iii) need to know when the market was competitive or collusive

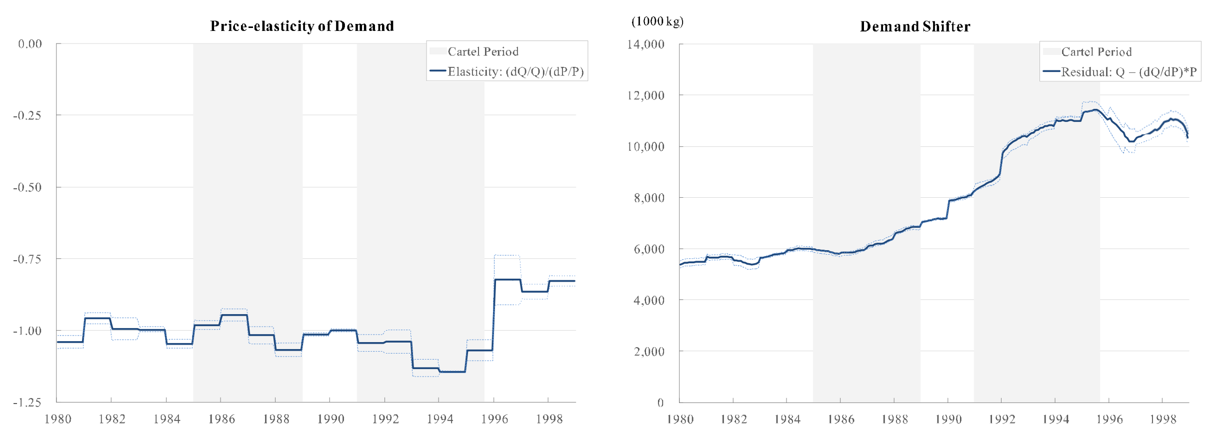
[33] See section 4.2 for the data considerations regarding outputs and capacities.

[34] To be precise, the identification of dP/dQ and the price elasticity of demand does not depend on the modeling choice between linear and log-linear functional forms. By contrast, the effective demand shifter, $\tilde{X}_t$, is calculated as the residual and depends on the exact definition of demand.

and (iv) need to model the supply side. We are willing to make this trade-off in the current empirical context because (i') the multicollinearity issue leads to hyper-sensitive demand estimates, (ii') a precise measure of consumers' perceptions do not exist, (iii') we know exactly when the firms were competing or colluding, and (iv') our objective (measuring the incentive to collude) necessitates a model of the supply side anyway. In other words, we are simply recycling parts of the data and the model to avoid making additional assumptions.[35]

Figure 4 (left) shows the price elasticity of demand was relatively stable at around -1, which is reasonable especially in the cartel period because a monopolist would not operate in an inelastic part of the demand curve. The right panel shows the effective demand shifter, $\tilde{X}_t$. As various industry experts noted, it exhibits a steady increase for most of the sample period, but the growth seemed to slow down in the late 1990s, which could be a factor for the cartel's collapse.

Figure 4: Demand Estimates for Vitamin C



Note: The dashed lines reflect the standard deviation of the estimated price coefficient in each year. The data frequency is monthly for prices and annual for outputs and costs (see section 4.2); hence, we first calculate the implied price coefficient at the monthly frequency and then take its annual average, interpreting its within-year variation as a reflection of measurement errors in outputs and costs. The standard deviation is relatively high in 1996 because the vitamin C price changed the most in that year.

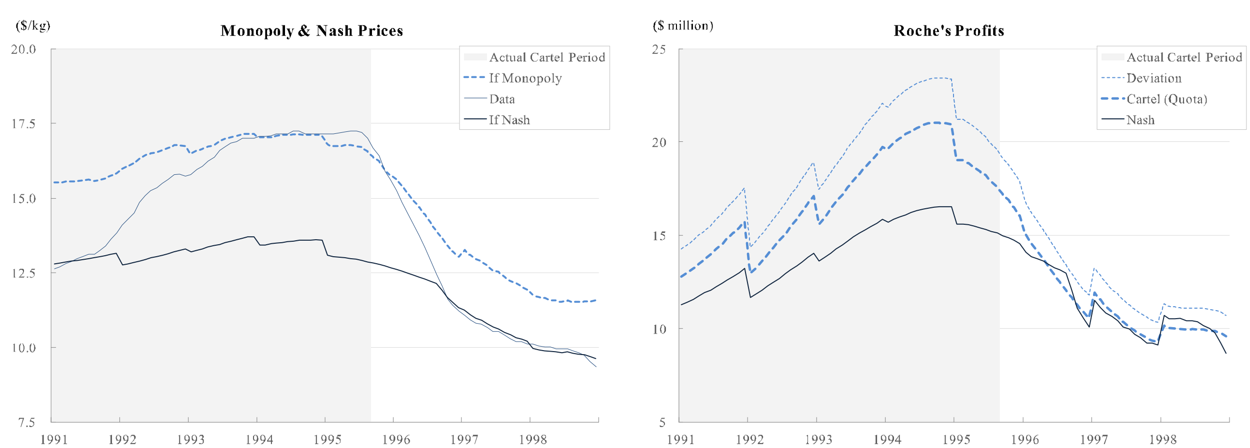
[35] To our knowledge, this identification strategy has not appeared in the academic literature on demand estimation, although Genesove and Mullin (1998) is a closely related example, in which the authors use cost data to compare direct and estimated measures of market power. Nevertheless, this approach is actually a standard practice in the cartel enforcement, according to Nathan Miller of Georgetown University. The reason is that antitrust agencies are in the same data environment as ours. Typically, they have access to firms' internal cost data, and they have evidence on the competitive/collusive regime. By contrast, perfect IVs for demand estimation are rarely available, and a complete list of demand shifters would be incredibly long. Our study heavily relies on the data from cartel enforcement; hence, it is not a coincidence that our data environment resembles that of antitrust practitioners.

5.3 Profits under Cartel, Static Nash, and Unilateral Deviation

Having estimated the demand and costs, we may calculate period profits, $\pi_{i,\tau|t}$, as in equation (1), for three different cases (recall from section 2 that our notation distinguishes between current period t and future period τ in expectation). The first is π^C–what a cartel firm would make given the demand, the fringe supply, and the quotas (we temporarily suppress subscripts for expositional simplicity). The second is π^D–what a cartel firm would make in the short run if it cheats (i.e., unilateral deviation from its quota while the other members are complying with their quotas). The third is π^N–the static Nash profit.

Figure 5 (left) plots three prices at each t (with $\tau = t$). The top line is the theoretical monopoly price (given the demand and fringe supply) that underlies the calculation of π^C; the bottom line is the Nash price that underlies π^N; and the middle line is the historical data. By contrast, the fact that the actual price converges to the monopoly price is surprising because no part of our model imposes restrictions on how the actual cartel price behaves in the data.[36] Two possibilities exist in principle. One is that our calculation of the theoretical monopoly price is wrong and the cartel failed to maximize its collective profit in reality. The other possibility is that our theoretical monopoly price is correct and the cartel did maximize its profit in reality. We are inclined to adopt the latter interpretation for two reasons.

Figure 5: Prices and Profits under Cartel and Static Nash



[36] One may wonder why the actual price took more than a year to achieve the monopoly level. One reason is that the cartel is illegal. Another reason is that the cartel needed to overcome the buyers' resistance by providing plausible explanations such as an increase of input costs, accidents at manufacturing facilities, and exchange-rate fluctuations. Yet another reason is that they had to coordinate and take turns in announcing price increases (see Marshall, Marx, and Raiff [2008]). All of these institutional contexts suggest the cartel could raise prices only gradually. See section 6.3 for the cartel pricing below the monopoly level.

First, the fact that Roche kept perfect data on transaction-level pricing and plant-level costs, as well as key information on its main rivals, suggests the company was in a position to precisely orchestrate the cartel pricing. BASF, its close partner, also kept a meticulous record of the country-level progress report on allotted quotas and actual sales by firm and by vitamin according to EC (2003). Second, our results for the three other markets (i.e., vitamin A, vitamin E, and beta carotene) indicate the same pattern. That is, the actual prices in these three markets also converge to their respective profit-maximizing levels, which suggests the fit is not a coincidence. Thus, we believe our assumptions and estimates are reasonable approximations to the actual developments in these vitamin markets and would serve a useful empirical model for the measurement of collusive incentives.

The vitamin C cartel decided to abandon its quota scheme on August 24, 1995, and ceased to exist, but the historical price data took approximately a year to reflect the full extent of downward adjustments. Recall that the price data aggregate all transactions at Roche, and typical contracts spanned three to 12 months. For this reason, Bernheim (2002a) allows a buffer period of 12 months since the cartel's collapse and then uses the subsequent period as the subsample representing a non-collusive era. We follow his data handling.[37]

Figure 5 (right) plots the corresponding period profits at Roche. The middle line is its cartel profits. It can make more profits for a short period of time if it oversells its quota (top line), but then it will have to live with Nash profits afterward (bottom line). Note the top line is always above the middle line (i.e., $\pi^D > \pi^C$ for all $t = \tau$), but the middle line decreased faster than the bottom line, and these two became indistinguishable (i.e., $\pi^C \approx \pi^N$) in the late 1990s. The reason is that Roche had a competitive cost structure but its quota did not reflect its full strength.[38] Thus, Roche's incentive compatibility constraint turns out to be the most binding.

5.4 Expected Values and Incentive Compatibility

Incentive Compatibility Constraints

These period profits serve as building blocks for constructing the payoffs from cooperation and defection in the form of expected present values, V^C and V^D, temporarily suppressing subscripts for expositional purposes. They correspond to equations (2) and (3) in section 2.

[37] Long-term (12-month) contracts do not alter our equilibrium analysis (see section 6.6).

[38] As the leader and the organizer of the cartels, Roche gave away fractions of its own "entitled" quotas (i.e., historical market shares in competitive periods) to its rival, so that it could sweeten the deal and finalize the organizational stage of collusion quickly. Unfortunately, no systematic record exists on these negotiations, but our merger analysis in section 7.2 carefully considers quota allocations.

In words, if a firm keeps colluding, it will earn π^C every period. If it cheats, it will earn π^D for three months, but then it will have to live with π^N forever.

Roughly speaking, we can compare V^C and V^D to see if the cartel is incentive compatible (i.e., $\Delta V_{i,t} \equiv V_{i,t}^C - V_{i,t}^D > 0$). Theoretically speaking, however, satisfying $\Delta V_{i,t} > 0$ in current period t is insufficient. If someone's $\Delta V_{i,\tau|t} < 0$ in some future period τ in expectation (given the available information at t), the problem becomes an essentially finitely repeated game and the collusive incentive evaporates immediately at t, not τ. Thus, the relevant condition for the cartel to be incentive compatible at t is that the *minimum* of $\Delta V_{i,\tau|t}$ be positive,

$$\Delta \underline{V}_{i,t} \equiv \min_{\tau \geq t} \Delta V_{i,\tau|t} > 0, \tag{8}$$

for all $i \in I$ at time t. This is the intuition behind the IC constraint (equation 4).

Cartel's Expectation about Future Demand and Fringe Supply

Given the prominent role of expectations and available information at each point in time, we gathered direct evidence on the cartel members' perceptions about future demand and supply. For the demand side, practically all industry experts agreed the demand for each vitamin was steadily growing over time, and that a decade was the relevant time horizon for planning purposes regarding major investments in facilities. Thus, we fit time trends (polynomials) to our estimates of $\tilde{X}_t$ and $\alpha_{1,t}$, and assume the firms rationally expected these fitted trajectories.[39] For the supply side, the cartel members' cost structures were common knowledge among themselves. Roche possessed detailed information on the rival firms' technologies and facilities. Production technologies had already matured by 1980. Hence, we assume firms rationally expected the virtually flat trajectories of marginal costs in the data and our estimates.

The most intriguing part of the vitamin C market is the role of the competitive fringe in China, which was negligible until 1991 but grew exponentially between 1992 and 1996. As we summarized in section 3.3, the emergence of China as a major exporter of vitamin C had been driven by a series of idiosyncratic events, including a major innovation in the mid 1980s, geopolitical shifts in the early 1990s, and the government's promotion in the mid 1990s. None of these events seemed to be correctly foreseen before 1992. EC (2003) reveals that even after three consecutive years of rapidly growing exports from China, the cartel was still forming static expectations as far as China was concerned. Our baseline model

[39] See section 6.4 for an alternative specification using adaptive expectations.

incorporates this evidence. Moreover, even if the cartel members were in fact fully rational, they would still engage in wishful thinking as long as a small probability existed that someone at the meeting really believed China would stop growing at some point. See section 2.3 for a thorough theoretical explanation of strategic wishful thinking: smart colluders fake dumb types.[40]

Results for Vitamin C

Figure 6 reports our estimates of the IC constraint for Roche, the most binding of the cartel members' conditions, in three steps. Panel A plots its $\Delta V_{\tau|t}$, the difference between the expected values of cooperation and defection in each future period (τ) conditional on the available information as of January each year (t). Each of the chain-looking lines represents a stream of $\Delta V_{\tau|t}$, and the colored dot marks the lowest point of each sequence, $\Delta \underline{V}_t \equiv \min_{\tau \geq t} \Delta V_{\tau|t}$, which needs to be positive for the cartel to be viable at t. Panel B collects these $\Delta \underline{V}_t$ for all t and shows the IC constraint was the tightest in 1995 and 1996, which is when the vitamin C cartel broke up in reality.

The discount factor, β, is difficult to identify. In Panels A and B, we set $\beta = 0.7$ (annual) for illustration purposes. Panel C plots $\Delta \underline{V}_t$ for a range of $\beta \in \{0.5, 0.6, 0.7, 0.8, 0.9\}$ to assess sensitivity, but the general patterns are similar and predict the violation of the IC constraint in 1995 or 1996. Our baseline $\beta = 0.7$ (or lower) perfectly rationalizes the breakdown of the cartel in 1995, although 0.8 or 0.9 will perform equally well if we allow for an additional parameter, such as $\phi > 0$, for some fixed cost of operating the cartel.

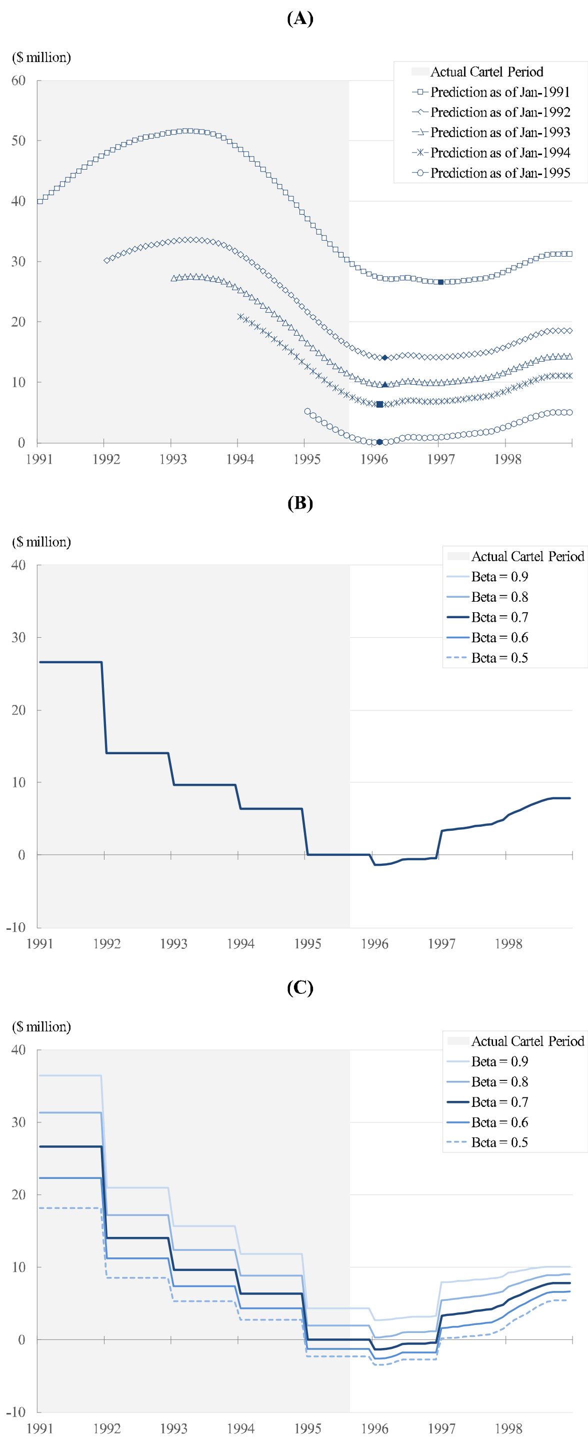
The reader might wonder whether $\beta = 0.7$, 0.8, or 0.9 is a reasonable level of patience for the large multinational companies in the chemicals/pharmaceutical sector. We believe so for two reasons. First, macroeconomic indicators such as real interest rates are not the most relevant factor, because the cartel was organized by the division managers in charge of vitamins, not the CEOs or the institutional shareholders of these firms.[41] The β in our model represents the personal time horizon of these individual managers concerning bonuses, promotions, and retirement.

Second, our β incorporates the perceived probability of an "exogenous" death of the cartel as well. The FBI's criminal investigation is one such example. An internal investigation by the firm's compliance office (or for due diligence prior to corporate transactions) is another.

[40] See section 6.1 for an alternative model with rational expectations and endogenous fringe.

[41] Recall that the heads of the vitamins divisions of Roche and BASF, as well as that of RP's Animal Nutrition division, started the cartels (section 3.2). The implicated managers kept the agreements secret from the top management and other divisions.

Figure 6: Roche's Incentive to Collude



Note: All values are multiplied by $1-\beta$ (i.e., rescaled as the average period profits) for expositional purposes.

An unexpected emergence of a substitute product or a competitive fringe could be yet another cause of exogenous terminations. These events did happen. Thus, our β, which embodies the subjective discount factor of individual managers (such as Dr. Kuno Sommer, then president of Roche's Vitamins and Fine Chemicals division), should be significantly lower than that implied by the risk-free rates. Dr. Sommer and other executives eventually went to American jail: a risky enterprise.

Vitamin A, Vitamin E, and Beta Carotene

In stark contrast with the vitamin C results, the same analysis yields completely different patterns for vitamins A, vitamin E, and beta carotene. In the Appendix, we find $\Delta \underline{V}_{i,t} > 0$ for Roche and all other colluders throughout the 1990s, suggesting no one's IC constraint was binding in these markets. Indeed, these three cartels were fully operational until the end, that is, until Rhône-Poulenc cooperated with the FBI and recorded the last cartel meetings in February 1999 (see section 3.2). Thus, our simple model explains the life and death of all four cartels for which reliable data exist. Section 7 further elaborates on the collapse of the vitamin C cartel, but first we pause to check the robustness of our findings.

6 Robustness

We use as much direct evidence as possible to inform our modeling choice and minimize the degree of arbitrariness in empirical analysis. Nevertheless, the following six aspects of the market lead us to consider alternative modeling assumptions.

6.1 Endogenous Fringe

China seemed the largest disturbing factor against the cartel stability. Our baseline model treated the Chinese government and SOEs as exogenous fringe players who built new plants for geopolitical reasons and dumped all domestic surpluses on the international market. We do not know of any generally accepted model of the Chinese government, and we believe the timing and the magnitude of their outputs reflected more idiosyncratic historical events than the usual, classical theory of firm behavior in market-based economies. Nevertheless, we may experiment with a model of China as "endogenous" fringe that responds to the international price. In Appendix C.1, we specify and estimate fringe supply as a function of lagged price. Moreover, we also consider the theoretical possibility that the cartel had perfectly known the fringe supply curve and had rationally played against it. We ask whether the cartel could have kept the price somewhat lower than the actual/monopoly level, so that fringe output would not expand as much as it actually did. Our analysis suggests such a "limit price" would have been too low for the cartel to be incentive compatible.

6.2 Renegotiating Quotas

We were relying on direct evidence from EC (2003) when we specified the cartel's sales quotas to be based on its members' pre-cartel market shares under oligopolistic competition. The record indicates that whenever a member proposed a revision of pre-existing quotas, the "Big Three" vehemently opposed the idea because they believed negotiations over quotas would never end and would destabilize cartel operations. Thus, we believe their use of historical market shares was a sensible choice given the practical costs of constant renegotiations. These pragmatic considerations notwithstanding, we may still ask whether the vitamin C cartel could have optimally renegotiated quotas to avoid its collapse in 1995. Appendix C.2 shows no quota reallocation could have saved the cartel.

6.3 Cartel Price below Monopoly Level

Section 5.3 showed the cartel pricing in the data converged to theoretical monopoly levels in all four vitamin markets within a year or two. Hence, our subsequent analysis treated cartel prices and monopoly prices interchangeably. However, existing studies suggest real-world cartels do not always entail monopoly prices. Appendix C.3 shows alternative prices below the monopoly level could not have relaxed the IC constraint in any significant way.

6.4 Adaptive Expectations

Our baseline model specifies the cartel members' beliefs in the following manner: (i) rational expectations on the demand side, (ii) rational expectations on the cartel firms' technology (i.e., cost structure), and (iii) static expectations on China (i.e., fringe supply). Each of these assumptions is based on direct evidence from UKCC (2001), Bernheim (2002a), and EC (2003). Nevertheless, the reader might wonder how the results would change if we impose informational frictions regarding (i) as well. All expert reports agree on a steady growth trend, but the demand could fluctuate around that trend. Adaptive expectations are a useful alternative, in which firms update predictions based on currently available information (i.e., past observations). Appendix C.4 uses adaptive expectations on $\tilde{X}_t$. We find this specification sharpens our model's prediction because it amplifies the impact of the slowdown of demand growth in 1995 and 1996.

6.5 Log Linear Demand

Section 5.2 showed our baseline demand analysis using the linear specification. We report the results using the log-linear specification in Appendix C.5. Neither the identification strategy nor the estimates of the price coefficient/elasticity changes significantly, but the implied demand shifter exhibits patterns that are inconsistent with the industry experts' testimony.

6.6 Long-term Contracts

As section 5.3 explained, a typical contract lasts for three to 12 months. Our baseline model assumes monthly spot-market transactions and abstracts from this multi-period aspect. In fact, explicitly incorporating such long-term contracts does not alter our analysis. In Appendix C.6, we show the IC constraint (4) remains the same even under an extreme assumption that all contracts last for 12 periods.

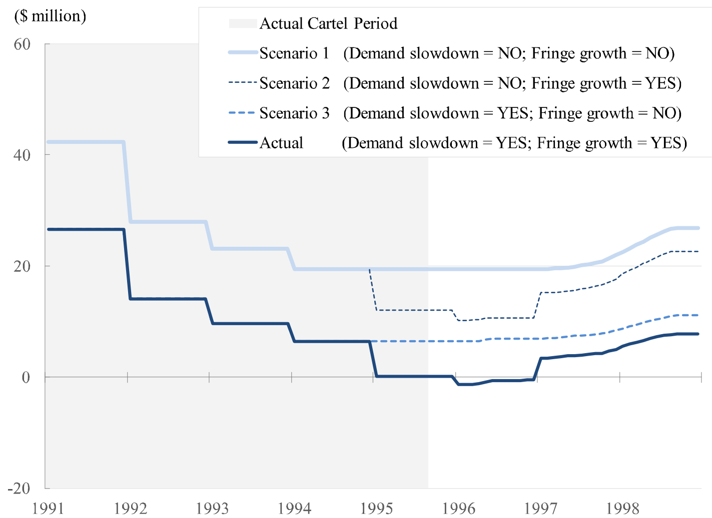
7 Counterfactual Simulations

Having estimated the model of endogenous cartel breakdown and checked its robustness, we may now answer our main questions. First, who killed the vitamin C cartel? Second, would an earlier consummation of the BASF-Takeda merger have helped prolong its life?

7.1 The Collapse of the Vitamin C Cartel and Its Causes

Figure 7 illustrates the effects of demand growth and fringe supply on the cartel's stability by showing Roche's IC constraint ($\Delta \underline{V}_t$) in four different scenarios. Scenario 1 at the top is the "dream world" counterfactual for the cartel, in which the demand shifter had not declined after 1994 (i.e., $\tilde{X}^{CF1}_{t>1994} = \tilde{X}^{Actual}_{1994}$) and the Chinese exports had stopped growing after 1994 (i.e., $Q^{CF1}_{fri,t>1994} = Q^{Actual}_{fri,1994}$), where the superscripts denote "counterfactual scenario 1" and "actual." Under these favorable conditions, Roche's incentive in 1995 would have been greater than that in reality as of 1992, 1993, or 1994 (i.e., $\Delta \underline{V}^{CF1}_{1995} > \Delta \underline{V}^{Actual}_{1992}$), which suggests the cartel could have survived throughout the sample period.

Figure 7: Effects of Demand and Supply



Note: The graphs depict Roche's IC constraint with $\beta = 0.7$. The relative positions of the four lines do not depend on the discount factor in any major way. See text for the three counterfactual scenarios.

Scenario 2 is the same as Scenario 1 on the demand side but incorporates the actual path of the Chinese fringe supply (i.e., $\tilde{X}^{CF2}_{t>1994} = \tilde{X}^{Actual}_{1994}$, and $Q^{CF2}_{fri,t} = Q^{Actual}_{fri,t}$ for all t). Hence, the difference between these two cases represents the impact of the Chinese output growth between 1994 and 1995, that is, what appeared to be the *coup de grâce* in the data. The simulated incentive in 1995, however, is still greater than the actual one in 1993 or 1994 (i.e., $\Delta\underline{V}^{CF2}_{1995} > \Delta\underline{V}^{Actual}_{1993}$), suggesting the survival of the cartel. Thus, despite the rapid expansion of the competitive fringe, the cartel could have survived beyond 1995 if the market size had been (weakly) increasing over time. Demand growth could play an important role.

Scenario 3 reaffirms the importance of demand by incorporating its actual path of downturn after 1994 while ignoring the Chinese supply growth after 1994 (i.e., $\tilde{X}^{CF3}_{t} = \tilde{X}^{Actual}_{t}$ for all t, and $Q^{CF3}_{fri,t>1994} = Q^{Actual}_{fri,1994}$). The difference between scenarios 1 and 3 reflects the impact of the sluggish demand after 1994. Roche's incentive in 1995 is still positive but as low as its actual level in 1994 (i.e., $\Delta\underline{V}^{CF3}_{1995} \approx \Delta\underline{V}^{Actual}_{1994}$).

Finally, the bottom line represents the reality, incorporating both the slowdown of the demand growth and the final phase of the Chinese supply expansion in the mid 1990s. This graph is the same one as Panel B of Figure 6 (i.e., $\beta = 0.7$). Note that none of the comparative statements in the above depends on the exact level of the discount factor in any major way.

7.2 Merger Helps Cartel Stability

The preceding analysis illustrated the effects of relatively simple factors on the incentive to collude. This section proceeds to investigate the role of market structure by simulating a merger between two of the four cartel firms in the vitamin C market, BASF and Takeda. Our primary motivation is to quantify the "coordinated effect" of merger in a theory-based manner. In antitrust policy, coordinated effects refer to the increased probability that firms engage in (tacit) collusion, which is different from a cartel as a legal concept, but economic theory does not necessarily distinguish between tacit and explicit collusion. A useful way to reconcile these perspectives is to regard our analysis of explicit collusion as a simple benchmark and interpret its collusive performance as an upper bound of what tacit colluders might be able to achieve.[42]

Historical Context and Model Setup

In reality, Takeda sold all of its vitamin businesses to BASF in 2001, that is, after the prosecution. The public relations disaster due to the cartel scandal was one reason, but a more fundamental reason was that the global pharmaceutical industry was undergoing a period of restructuring in which conglomerates pulled out of non-core, low-margin businesses such as bulk chemicals (Yamaguchi 2000).[43] Instead of 2001, we simulate a similar merger in 1990, a year before the vitamin C cartel started.

We construct this counterfactual scenario as follows. First, Takeda's vitamin C facilities were larger and more efficient than BASF's, which was shut down after the actual merger in 2001. Thus, the merged entity should inherit Takeda's cost structure. Second, neither the long-run physical capacities nor the reported "nameplate" capacities were binding at any point during the sample period (see section 4.2). BASF's closure of its own facilities after the actual merger is indicative of excess capacity indeed. Hence, we do not simulate a situation in which capacities are binding.[44] Third, the actual cartel allocated quotas based on the historical market shares before the cartel, that is, competitive outcomes before 1991. We construct a counterfactual version of such historical market shares by re-computing hypothetical Nash

[42] See Awaya and Krishna (2016) for a theoretical work to distinguish between tacit and explicit collusion. Kaplow (2013) elaborates on exactly what constitutes price fixing.

[43] We thank Hidemaru Yamaguchi, health care and pharmaceutical research analyst at Citigroup Global Markets Japan, for sharing his knowledge on these corporate transactions and industry history.

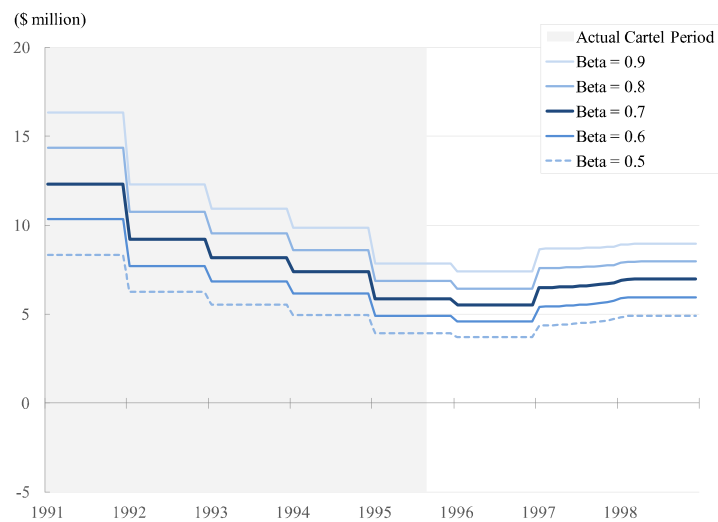
[44] Recall from section 4.2 that the appropriate notion of "capacity" for the Cournot model in our context (à la Kreps and Scheinkman 1983) is that of quantity precommitment in the form of short-run production plans at the monthly frequency, which should be distinguished from either the long-run physical capacity (which was never binding) or the reported nameplate capacity (which was never binding either).

outcomes in 1990 between three firms (Roche, E. Merck, and the consolidated BASF-Takeda) instead of four in reality.[45] Under this alternative setup, we recalculate streams of profits under the cartel, unilateral deviation, and static Nash.

Results

Figure 8 draws the counterfactual $\Delta \underline{V}_{roche,t}$ in which the BASF-Takeda merger had consummated before the beginning of the vitamin C cartel in 1991. Regardless of the level of $\beta \in [0.5, 0.9]$, Roche's IC constraint is relaxed and positive throughout the cartel period, because Roche's quota becomes considerably larger after the four-to-three merger. Under this alternative market structure, the vitamin C cartel would have continued operations at least until February 1999 (i.e., when the FBI and the DOJ intervened).

Figure 8: If BASF Had Acquired Takeda's Vitamin C Business by 1991



To illustrate the extent of "coordinated effect" (in our particular sense), Table 3 summarizes the welfare outcomes as of 1998. The average price was $9.8 in reality, which would have increased by 1.7% under a hypothetical Cournot competition among the three firms instead of four. This small change reflects the merger's "unilateral effect" (i.e., a higher markup under oligopolistic competition among fewer firms). However, if we account for the possibility that the merger could have materially improved the cartel's incentive compatibil-

[45] The reader might recall from sections 3.2 and 5.4 that Roche ceded some of its "historically entitled" quota to Takeda and BASF at the beginning of the vitamin C cartel. See the Appendix for details and a robustness check.

ity, the prevailing 1998 price would have been 18.1% higher at \$11.58 and consumer surplus 68.6% lower.[46]

Table 3: Welfare Analysis with Coordinated Effects

	No merger (four firms)	Merger simulation (three firms)	
Unilateral effect	No	Yes	Yes
Coordinated effect	No	No	Yes
Price (\$/kg)	9.81	9.98	11.58
	(±0%)	(+1.7%)	(+18.1%)
Output (1000MT)	70.5	69.5	37.9
	(±0%)	(−1.4%)	(−46.3%)
Consumer surplus (\$million)	410.3	398.7	186.7
	(±0%)	(−2.8%)	(−68.6%)

Note: Annualized average outcomes in 1998.

Decomposing the Effect of Merger on the IC Constraint

How does the merger change the incentive to collude? We first clarify the conceptual framework to answer this question, and then explain our empirical results. Let us temporarily ignore the "minimum" part of the IC constraint in equation (4). We can rearrange the original terms from equations (2) and (3) to decompose the IC constraint as follows,

$$\underbrace{\sum_{s\geq\tau+3}\beta^{s-1}\pi^{C}_{i,s|t}}_{\text{on-path continuation value}} - \underbrace{\sum_{s\geq\tau+3}\beta^{s-1}\pi^{N}_{i,s|t}}_{\text{punishment continuation value}} \geq \underbrace{\sum_{s=\tau}^{\tau+2}\beta^{s-1}\pi^{D}_{i,s|t}}_{\text{(gross) deviation gain}} - \underbrace{\sum_{s=\tau}^{\tau+2}\beta^{s-1}\pi^{C}_{i,s|t}}_{\text{forgone on-path gain}}. \quad (9)$$

Relabeling the four terms as $V^1_{i,\tau|t} - V^2_{i,\tau|t} \geq V^3_{i,\tau|t} - V^4_{i,\tau|t}$ for notational convenience, the IC requires that the difference between the on-path continuation value ($V^1_{i,\tau|t}$) and the value under punishment ($V^2_{i,\tau|t}$) be greater than the net deviation gain in the short run ($V^3_{i,\tau|t} - V^4_{i,\tau|t}$).

A merger tends to increase both $V^1_{i,\tau|t}$ and $V^2_{i,\tau|t}$ because fewer firms split the market, either cooperatively ($V^1_{i,\tau|t}$) or non-cooperatively ($V^2_{i,\tau|t}$). In addition, $V^3_{i,\tau|t}$ also increases, because a deviation brings a bigger market share with fewer firms. However, the rate of the increase differs among $V^1_{i,\tau|t}$, $V^2_{i,\tau|t}$, and $V^3_{i,\tau|t}$. Hence, even with symmetric cost and an equal split of the market, the left-hand side minus the right-hand side of (9) may change

[46] If we further incorporate the possibility that China might have behaved differently from the reality, the result will be slightly less dramatic, with the counterfactual price at \$11.28 (15.1% higher than the baseline). See the Appendix for details, in which we construct a "worst-case scenario" based on China's highest historical output in 1996.

non-monotonely with the number of firms. Consequently, the direction and the magnitude of a merger's effect on the IC constraint is theoretically ambiguous (see Appendix D.1).

Moreover, the true IC constraint must be evaluated at its minimum across firms and time. The identity of the firm with the most binding IC constraint (i.e., Roche) does not change in the case of vitamin C, but both demand and costs change over time, and hence the crucial period τ that entails $\min_{i \in I, \tau \geq t} \left(\Delta V_{i,\tau|t} \right)$ could also change. For these reasons, the assessment of the merger's impact on cartel stability is a theoretical as well as empirical problem.

Table 4: Accounting for Changes in the IC Constraint

($ million)	No merger (1)	Merger (2)	Change (3) = (2) − (1)	Contribution to IC (4)
Period τ that minimizes $\Delta V_{\tau \mid Aug\text{-}'95}$	Feb-1996	Dec-1996		
On-path cont. value (V^1)	93.3	97.9	+4.5	78.6%
Punishment value (V^2)	91.8	91.2	−0.6	10.4%
Gross deviation gain (V^3)	14.3	11.7	−2.6	45.5%
Short-run on-path gain (V^4)	12.8	10.9	−2.0	−34.5%
Net on-path cont. value ($V^1 - V^2$)	1.6	6.7	+5.1	89.0%
Net deviation gain ($V^3 - V^4$)	1.5	0.8	−0.6	11.0%
IC constraint: $\left(V^1 - V^2\right) - \left(V^3 - V^4\right)$	0.1*	5.9	+5.8	100.0%

Note: The IC constraint and its components as of August 1995 with $\beta = 0.7$. Note the baseline IC constraint is not exactly zero (0.1), but our narrative in the main text ignores this small numerical difference.

Let us investigate the IC constraint and its components in August 1995, the month in which the actual cartel met for the last time and abandoned the agreement. Table 4 compares the four components of $\Delta \underline{V}_{Aug\text{-}'95}$ in the baseline model (column 1) and the merger simulation (column 2) when $\beta = 0.7$. Their differences appear in column 3, and column 4 decomposes the net change in $\Delta \underline{V}_{Aug\text{-}'95}$ into the percentage contribution of each factor.

The increase in $V^1_{i,\tau|t}$ accounts for the lion's share, followed by the decreases in $V^3_{i,\tau|t}$ and $V^2_{i,\tau|t}$. The contribution of $V^4_{i,\tau|t}$ is negative and does not align with that of $V^1_{i,\tau|t}$, but note the most binding τ (in expectation as of August 1995) has changed from February 1996 (baseline) to December 1996 (counterfactual), so $V^1_{i,\tau|t}$ and $V^4_{i,\tau|t}$ have no reason to move in the same direction.

Thus, our IC accounting shows the hypothetical BASF-Takeda merger would have prolonged the life of the vitamin C cartel primarily by increasing the on-path continuation value for Roche, whose quota would have been sufficiently generous under the three-firm oligopoly. This analysis demonstrates the benefits of a theory-based empirical accounting for a better understanding of the relationship between mergers and collusion.

8 Conclusion

Repeated-game theory is famous for its conceptual sophistication but notoriously difficult for empirical implementation. With suitable data and institutional details, however, the framework is useful for issues of practical importance. Specifically, it could (i) explain the life and death of the four major vitamin cartels, (ii) quantify the effects of demand and supply on the incentive to collude, and (iii) predict the "coordinated effects" of merger.

In the course of empirical analysis, we effectively tested one of the most fundamental predictions of repeated games–that cooperation breaks down when the incentive compatibility is violated. Our model does not distinguish between explicit and tacit collusion, and our empirical context does not involve private monitoring, but such features could be relevant in other data environment and might represent fruitful directions for future research. Our findings also suggest the possible interactions between cartel and merger enforcement for better antitrust policy.

Appendix (For Online Publication)

This Appendix contains four different kinds of information. The first is additional materials from the Bernheim reports and other data sources. The second is the empirical analysis of vitamin A, vitamin E, and beta carotene. The third and fourth are additional results concerning our main analysis and merger simulation, respectively. The following table lists the contents.

Table 5: Contents of the Appendix

Appendix	Contents	Corresponding main text
A	Additional facts and data	Sections 1, 2, 3, and 4
A.1	Top-10 antitrust violations	
A.2	Profile of the Chinese SOEs	
A.3	Prices of specific products by firm	
A.4	Inputs for vitamin production	
A.5	Unit cost and output	
B	Empirical analysis of vitamin A, vitamin E, and beta carotene	Section 5
C	Robustness checks	Section 6
C.1	Endogenous fringe supply	
C.2	Renegotiation of quotas	
C.3	Cartel price below monopoly level	
C.4	Adaptive expectations on future demand	
C.5	Log linear demand	
C.6	Long-term contracts	
D	Merger simulation	Section 7.2
D.1	Merger may increase or decrease collusive incentives	
D.2	Roche's quota concession	
D.3	Aggressive fringe supply	

Appendix A: Additional Facts and Data

A.1 Top-10 Antitrust Violations

In section 1, we said the international vitamin cartels are among the largest antitrust cases in history. Table 6 ranks the top-10 antitrust cases in America by the amount of fine. The DOJ has considered the car parts cartels the largest case, but Roche's fine (in the 1999 dollars) tops the list by firm, even without adjustment for inflation.

Table 6: Top-10 Sherman Act Violations and Corporate Fines

Rank	Product	Defendant	Country	Fiscal year	Geographic scope	Fine (million)
1	Vitamins	Roche	Switzerland	1999	International	\$500
1	LCD panels	AU Optronics	Taiwan	2012	International	\$500
3	Car parts	Yazaki	Japan	2012	International	\$470
4	Car parts	Bridgestone	Japan	2014	International	\$425
5	LCD panels	LG Display	Korea	2009	International	\$400
6	Air Transport	Air France & KLM	France & Netherlands	2008	International	\$350
7	Air Transport	Korean Air	Korea	2007	International	\$300
7	Air Transport	British Airways	UK	2007	International	\$300
7	DRAM	Samsung	Korea	2006	International	\$300
10	Vitamins	BASF	Germany	1999	International	\$225

Note: Not adjusted for inflation. Ranking as of September 12, 2016. *Source*: U.S. Department of Justice, Antitrust Division.

A.2 Profile of the Chinese SOEs

Four major firms appear on records in the 2000s (UKCC 2001, Bernheim 2008). First, North East Pharmaceutical is an SOE located in Shenyang, Liaoning Province, with its historical roots in Takeda's pre-war plant.

Second, Jiangsu Jiangshan Pharmaceutical was founded in 1990 as a joint venture of the provincial government and other entities in Jingjian, Jiangsu Province.

Third, Shijiazhang Pharmaceutical is another SOE controlled by the provincial government. It was established in 1997 as an amalgamation of four drug companies in Shijiazhuang, Hebei, and its vitamin C arm is Weisheng.

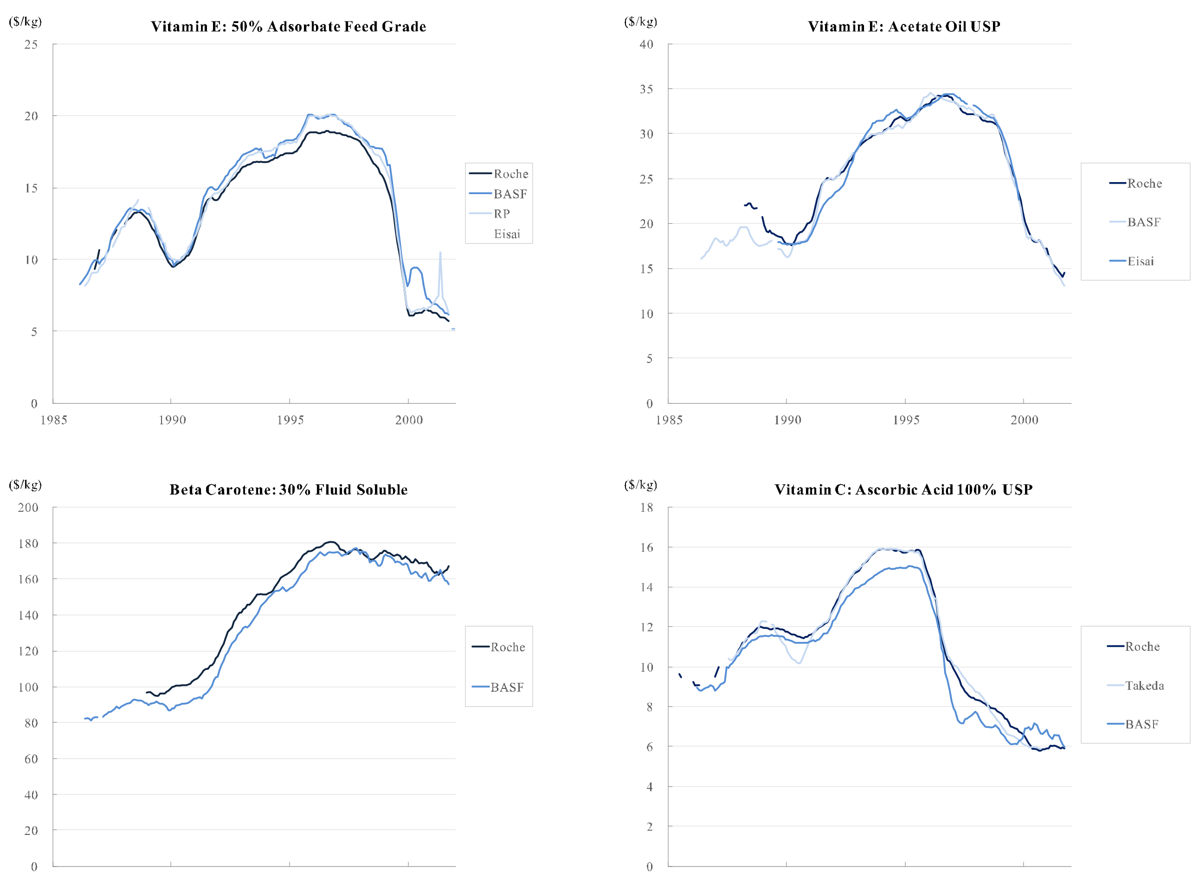
Fourth, North China Pharmaceutical is one of the largest drug makers in China and is also under the control of Hebei Province's State Asset Management Committee. Its vitamin C arm is Hebei Weierkang ("Welcome"). In 2004, exports accounted for only 0.03% of North China's revenue, which suggests its primary focus had been the domestic market.

Most of the private entrants in the mid 1990s seemed to have disappeared by the early 2000s, except for those that survived by becoming part of SOEs. SOEs are dominant players in the Chinese pharmaceutical industry because compliance with the quality and safety standards requires close connections to the government (Xiang, Zhang, Chen, and Watanabe 2007), and because the manufacturing processes of chemical products are typically capital intensive (capital markets in China are controlled by state-owned banks).

A.3 Prices of Specific Products by Firm

In section 3.1, we discussed the similarity of prices across firms and different grades. Figure 9 confirms this observation by showing firm-level prices of four specific vitamin products (Vitamin E 50% Adsorbate Feed Grad, Vitamin E Acetate Oil USP, Ascorbic Acid 100% USP, and Beta Carotene 30% Fluid Soluble), which are highly correlated and close to each other in levels.

Figure 9: Examples of Vitamin Prices by Firm and Grade



The reader might wonder whether the discrepancies across firms are small. Bernheim (2002a) and other expert reports suggest they are. Each firm was headquartered and produced in a different country and currency zone (note the European common currency was not effective during our sample period). The prices in the graphs are denominated in the US Dollar and do not account for exchange-rate fluctuations between the Swiss Franc, the Deutsche Mark, the French Franc, and the Japanese Yen. Moreover, firms used different accounting rules. Despite these random factors, the graphs exhibit discrepancies in the range of approximately 5% for over 15 years.

A.4 Inputs for Vitamin Production

In section 3.1, we discussed the production processes of vitamins. Table 7 lists raw materials and intermediates for each vitamin.

Table 7: Key Chemical Ingredients Required for Vitamin Synthesis

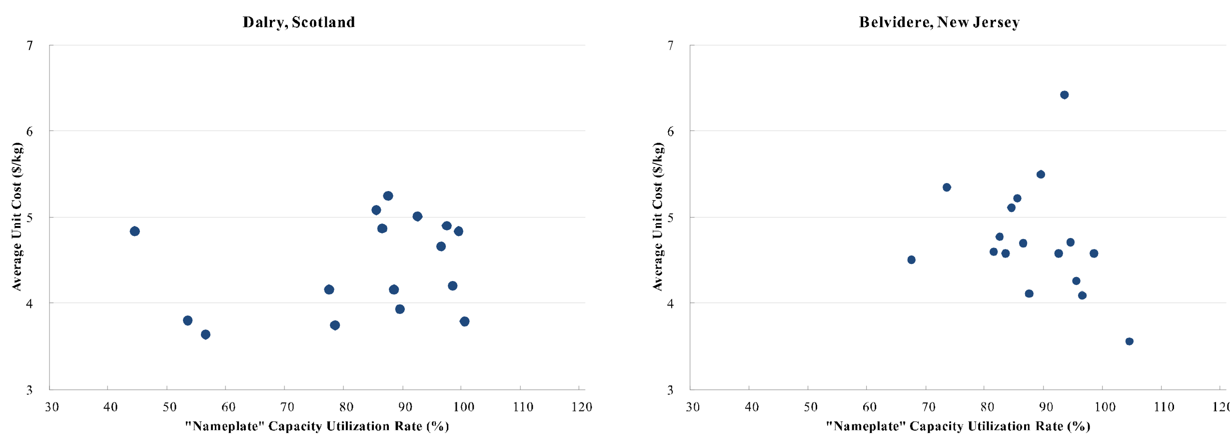
Vitamin	Intermediates	Raw Materials
A	Pseudoionone	Acetone, acetylene, isobutylene, butenediol, formaldahyde
B1	Grewe diamine	Ethylene, Prymidine, malononitrile, acronynitrile, carbon monoxide, cetamidine, butyroloctone, methyl acetate, hydrochloric acid, ammonia, carbon disulphate
B2, synthetic	Ribose	
B2, fermented		Sugars
B3	Methylglutaronitrile, beta picoline, 3-cryanopyridine, methylethylpyridine	Ethylene, nitric acid, farmaldyhyde, ammonia
B4	Trimethylamine	Hydrochloric acid, ethylene oxide
B5	Pantolactone, beta-alanine	Iso butryaldehyde, bydrogen yanide, hydrochloric acid, acrylonitrile, ammonia, caustic soda, calcium hydroxide
B6		Oxazole, dienophile
B9		Acetone or acrolein, chlorine gas, guanadine, cyanoethyl acetate, sodium ethoxide, nitric acid, hydrogen gas, glutamic acid, benzoic acid
B12, fermented		Sugars, nitrogen compounds
C	Sorbitol	Glucose
D3		Cholesterol
E, synthetic	Isophytol, trimethylhydroquinone	Acetone, acetylene, isobutylene, napha, formaldahyde
H	Thiolactone	Furnaric acid or diketene, cysteine, thiophene, phosgene gas
Beta carotene		Acetone, ocetyene, triphenylphosphine
Canthaxanthin	Beta carotene	15-carbon compounds
Astaxanthin	Canthaxanthin	15-carbon compounds

Source: Bernheim (2002a), Connor (2007, 2008).

A.5 Unit Cost and Output

Bernheim (2002b) uses Roche's internal cost data to investigate the relationship between the unit cost and the "nameplate" capacity utilization. Figure 10 plots Roche's variable cost curves at production facilities in two different locations, Scotland (Dalry) and New Jersey (Belvidere).[47] Three patterns deserve attention. First, the nameplate capacity utilization could exceed 100%. As we explained in section 4.2, the reported capacities correspond to neither of the two economically relevant notions of capacity (i.e., long-run physical limits and short-run production plans), with ad hoc definitions that vary across time and plants. We simply interpret these nameplate capacities in the graphs as a proxy for production here.

Figure 10: Average Variable Cost and Reported Capacity Utilization by Plant



Second, the cost structures are remarkably similar across these plants in two different countries. We have 16 annual observations for Dalry (mean \$4.43 and standard deviation \$0.55) and 17 for Belvidere (\$4.75 and \$0.65, respectively). The t-test does not reject the null hypothesis that the two means are identical at any conventional significance levels.

The third noteworthy feature is the flat slopes of the two cost curves. Linear regressions do not reject the null hypothesis that the slope coefficients equal zero at any conventional significance levels, either at the individual plant level or in the pooled sample. We take this as the evidence of constant returns to scale, which is consistent with our modeling assumption in section 2.[48]

[47] Roche retired another plant in Germany (Grenzach), whose data do not represent normal operations and therefore are not shown here.

[48] This observation does not necessarily conflict with the conventional notion that "increasing returns to scale" characterize the production technologies for bulk chemicals, because firms need to build large facilities to achieve these cost curves in the first place.

Appendix B: Empirical Analysis of Vitamin A, Vitamin E, and Beta Carotene

Our main text (section 5) focused on the vitamin C market because it was the only one (with useful data) in which the cartel collapsed without antitrust intervention. The other three cartels survived until the authority busted them in February 1999, and hence analyzing them might appear less interesting at first glance. Nevertheless, we have at least three reasons to study them.

First, their data patterns help us recognize the overall similarity across vitamin markets. Figure 11 shows the prices were significantly higher in the 1990s than in the 1980s despite virtually flat costs. The total outputs were steadily increasing because the world demand was growing for each vitamin (Figure 12 shows upward trends in our estimates of the effective demand shifter, $\tilde{X}_t$). These pictures are familiar from our analysis of the vitamin C market (i.e., Figures 3 and 4).

Second, a closer look at the output graphs reveals two intriguing differences across markets. One is the variation in market structure with regard to the cartel firms. Beta carotene is duopoly, vitamin A is triopoly, and vitamin E is quadropoly. The market structure of vitamin C was the closest to that of vitamin E in the sense that both involved the Big Three European firms and one Japanese firm. Another important difference is the size of the competitive fringe. None existed in beta carotene; vitamins A and E had some small suppliers. Vitamin C resembled the latter two markets until 1992, in which the Chinese SOEs massively increased their presence.[49]

Third, a comparison of the estimation results corroborates both the validity and the usefulness of our framework. Figure 5 in section 5.3 showed the actual cartel price converged to the monopoly level in the vitamin C market; the three left panels of Figure 13 show similar trajectories in the other markets.[50] The right panels draw Roche's IC constraint, $\Delta \underline{V}_t$, for a range of β. All lines are comfortably above zero, indicating the stability of these three cartels throughout the 1990s.

[49] Why were Chinese researchers successful in the process innovation for only vitamin C and not other vitamins? We have found no systematic evidence to answer this interesting question. In terms of dollar value, vitamin E was the largest market in the 1990s. In terms of profit margin and the softness of competition, beta carotene was the most lucrative market. Hence, relative market sizes cannot rationalize the innovation or the rapid growth of the SOEs' exports in the vitamin C market. See section 3.3 for broader historical contexts.

[50] The beta carotene price might appear to bounce between the two theoretical benchmarks. The difference between the monopoly price and the Nash price is relatively small because of its duopoly market structure.

Figure 11: Prices, Costs, and Quantity of Vitamin A, Vitamin E, and Beta Carotene

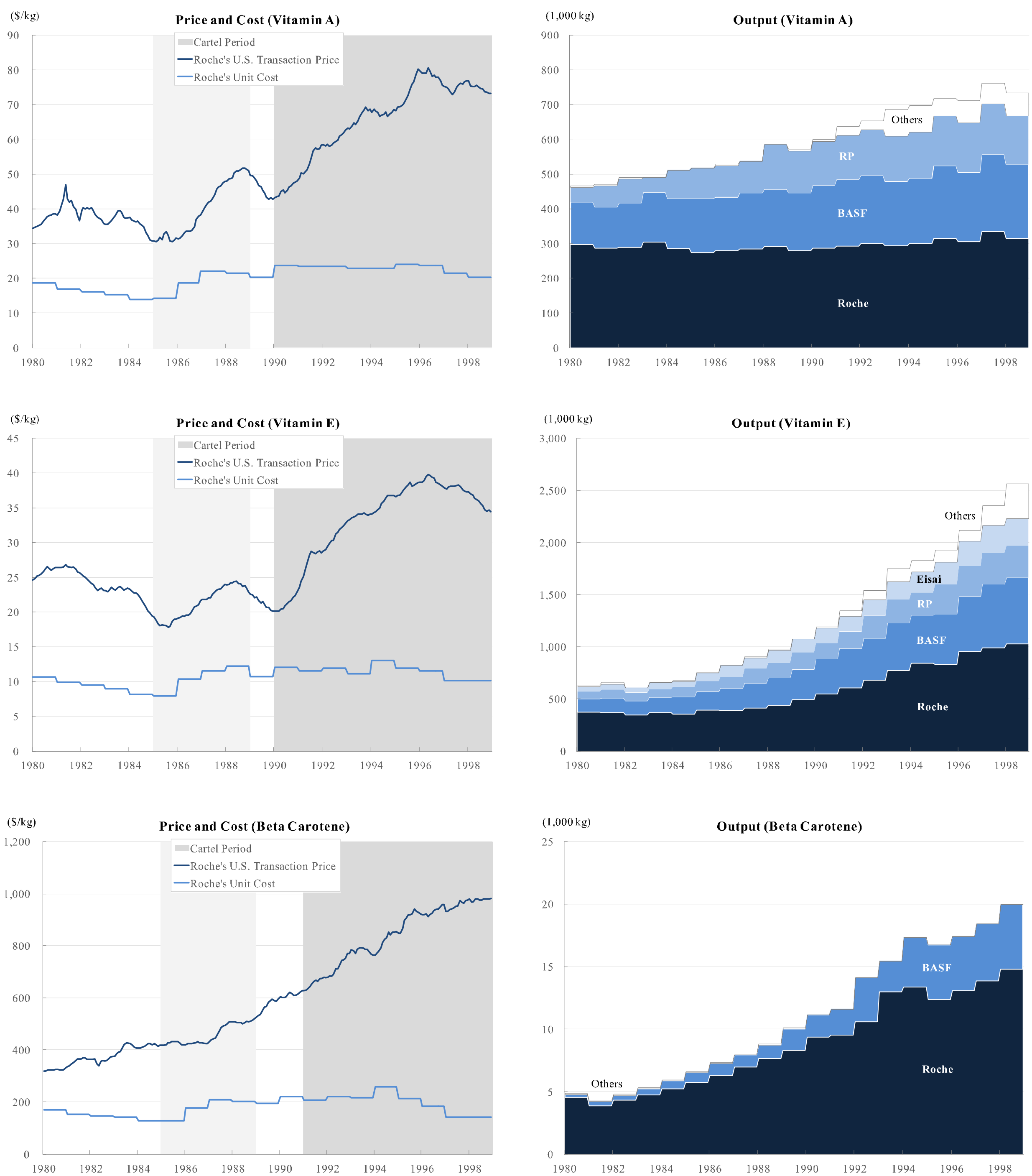


Figure 12: Demand Estimates for Vitamin A, Vitamin E, and Beta Carotene

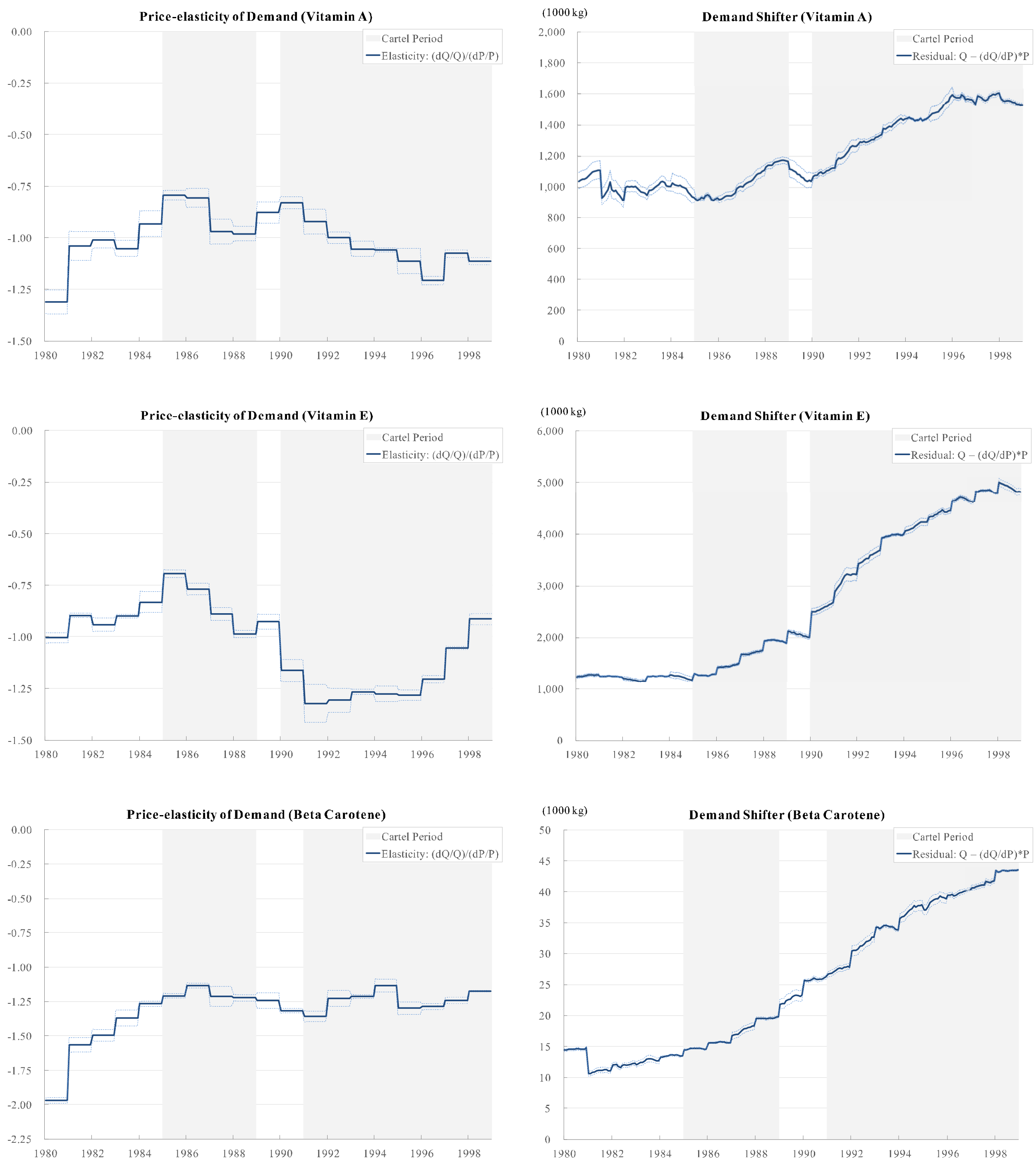
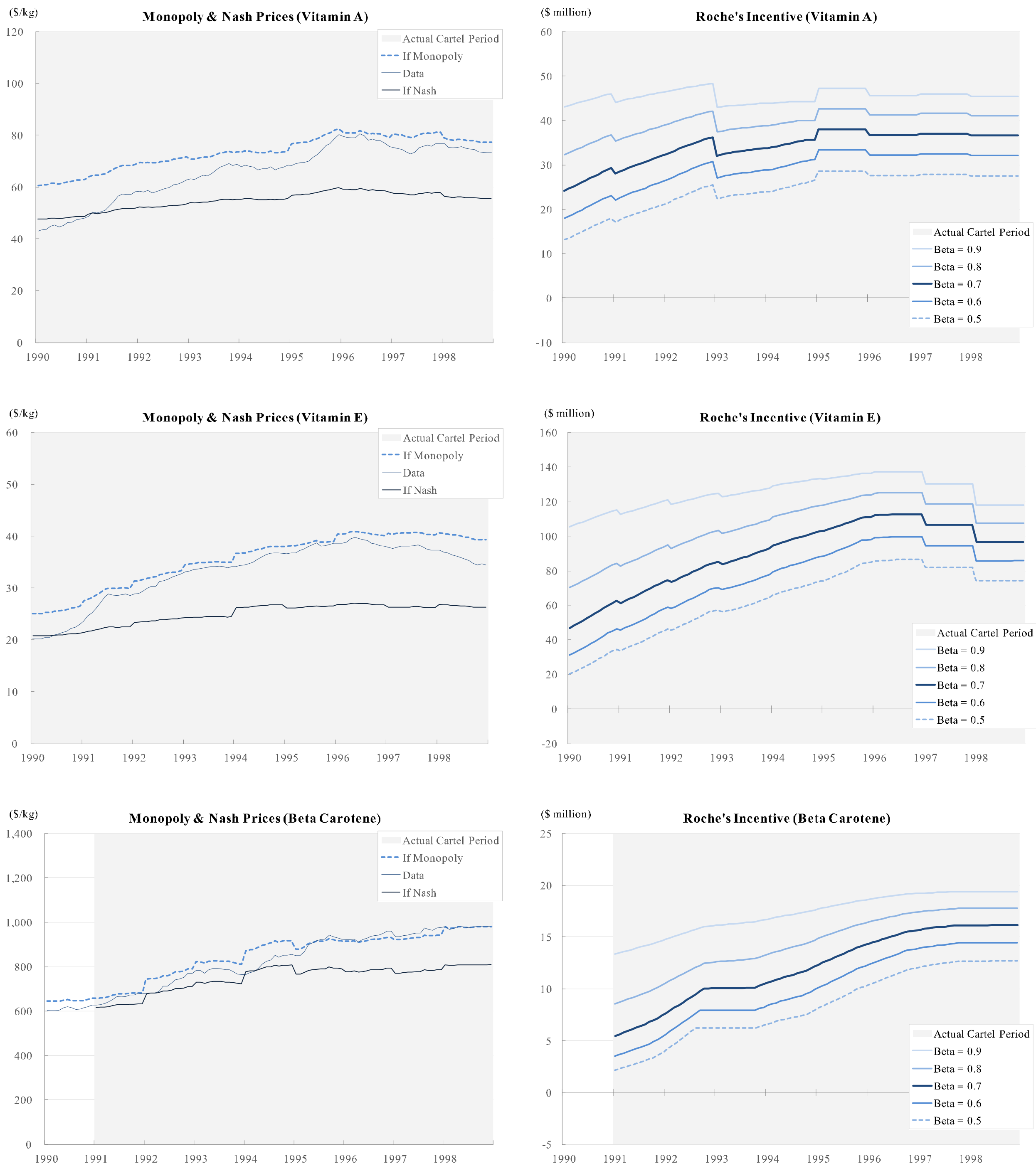


Figure 13: Collusion Incentives for Vitamin A, Vitamin E, and Beta Carotene



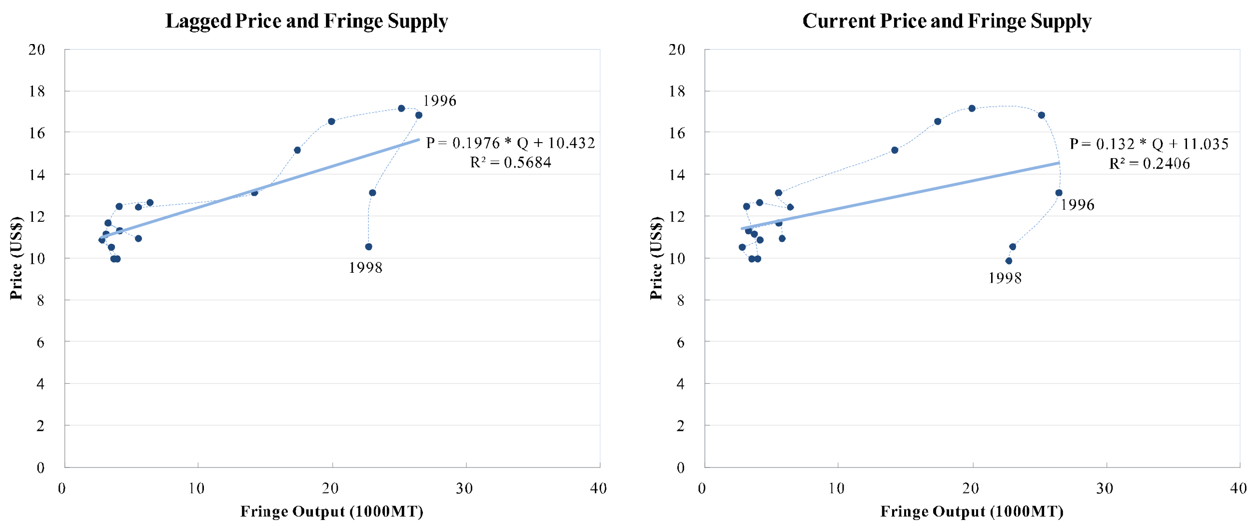
Appendix C: Robustness Checks

C.1 Endogenous Fringe (continued from section 6.1)

Data Patterns

Figure 14 shows two upward-sloping supply curves in the vitamin C market, plotting the total fringe output against the lagged price (left) and the concurrent price (right), respectively. The fringe output strongly correlates with the lagged price but only modestly with the current price, which is consistent with the institutional background of the Chinese suppliers' staggered entry and expansion in section 3.3.

Figure 14: Fringe Supply Curves



Our Question

Let us depart from our baseline model in which we characterized the Chinese SOEs as purely exogenous fringe. Suppose a lagged supply curve such as Figure 14 (left) adequately describes the *endogenous* supply response from China. With time, the cartel firms might be able to learn and estimate this fringe supply function. Eventually, they might even come up with a "better" cartel output level that strikes the right balance between raising the price to increase current profits and lowering the price to keep the Chinese exports low.

To investigate this possibility, we study the following situation *after* the sample period, at which point we may reasonably assume the (former) cartel firms had enough data observations to learn the fringe supply curve. Suppose the (former) cartel firms play a Markov perfect equilibrium (MPE) after the prosecution in 1999 (now that we assume the competitive fringe reacts to past prices, an MPE is the adequate non-collusive equilibrium concept corresponding to a static Nash in our baseline model). Suppose the price is at its stationary level.

We ask whether the firms could possibly agree on a new cartel scheme under this situation. Our answer is "no," as we explain in the following.

Model

To ensure our analysis is robust to those considerations, we consider the following model: we fix $\tilde{X}_t = x$, $\frac{dP}{dQ} = \psi$, and $c_{i,t} = c_i$ at the level of 1998. The model is the same as in section 2 except that firms rationally expect that the fringe supply depends on past K-period prices linearly:

$$Q_{fri,t} = Q_{fri}\left(P_{t-1}, ..., P_{t-K}\right) = a_{fri} + \sum_{k=1}^{K} b_{fri,k} P_{t-k},$$

where a_{fri} and $b_{fri,k}$ are coefficients of the fringe supply curve. The idea is to capture the notion that Chinese SOEs were building new plants and entering the international market for the first time, partially as an endogenous response (albeit delayed) to the high prices in the international vitamin C market.

The inverse demand function for the total output by the players/cartel participants, $Q = \sum_{i \in I} Q_i$,[51] is

$$P = -\psi \times (x - Q_{fri}) + \psi \times Q, \text{ or } Q = x - Q_{fri} + \frac{P}{\psi}. \tag{10}$$

Optimal Cartel Supply We first derive the optimal cartel supply, where the players jointly maximize the cartel profit. As in section 2, because Roche plays a leading role, the joint profit is calculated assuming each player has the same marginal cost as Roche, denoted by c.

Because the environment is stationary, the state variable for calculating the optimal cartel profit is $(P_{-k})_{k=1}^{K}$, where P_{-k} is the price k periods before. In particular, the value function $V((P_{-k})_{k=1}^{K})$ satisfies that the cartel picks Q (given (10), equivalent to picking P)

[51] Here we did not allow the Chinese firms to join the cartel, because even in the UKCC's report written in 2001, the incumbent firms in Europe and Japan characterized them as an aggressive and untamable fringe.

to maximize

$$V\left(P_{-1:-K}\right)=\max_{P}\left(P-c\right)\left[x-Q_{fri}+\frac{P}{\psi}\right]+\beta V\left(P,P_{-1:-K+1}\right), \tag{11}$$

where $P_{-l:-k}=(P_{-l},P_{-l-1},...,P_{-k})$. Note that Q_{fri} depends on $(P_{-k})_{k=1}^{K}$. The first-order condition is

$$x-Q_{fri}+\frac{2P-c}{\psi}+\beta V_1\left(P,P_{-1:-K+1}\right)=0,$$

where V_k is the derivative of the kth argument. By the envelope theorem, we have

$$V_k\left(P_{-1:-K}\right)=-\left(P-c\right)b_{fri,k}+\beta V_k\left(P,P_{-1:-K+1}\right)$$

and so

$$V_k\left(P_{+k-1:+1},P,P_{-1:-K+k}\right)=-\left(P_{+k}-c\right)b_{fri,k}+\beta V_{k+1}\left(P_{+k:+1},P,P_{-1:-K+k+1}\right),$$

where P_{+k} means the price k periods ahead. Hence, we have

$$\beta V_1\left(P,P_{-1:-K+1}\right)=-\sum_{k=1}^{K}\beta^k\left(P_{+k}-c\right)b_{fri,k}. \tag{12}$$

Given (12), the first-order condition can be written as

$$x-Q_{fri}+\frac{2P-c}{\psi}-\sum_{k=1}^{K}\beta^k\left(P_{+k}-c\right)b_{fri,k}=0,$$

or

$$P=\frac{|\psi|\left(x-a_{fri}-\sum_{k=1}^{K}b_{fri,k}P_{-k}\right)-|\psi|\sum_{k=1}^{K}\beta^k\left(P_{+k}-c\right)b_{fri,k}+c}{2}. \tag{13}$$

We prove a linear solution exists:

Lemma 1 *There exists a linear solution for (11):* $P=a_{cartel}+\sum_{k=1}^{K}b_{cartel,k}P_{-k}$.

Let $\bar{P}\left(\{P_{t-k}\}_{k=1}^{K}\right)$ and $\bar{Q}\left(\{P_{t-k}\}_{k=1}^{K}\right)$ denote the optimal price and total quantity given the past prices $\{P_{t-k}\}_{k=1}^{K}$.

Proof. With discounting, it suffices to guess the solution takes the form

$$a_{cartel} + \sum_{k=1}^{K} b_{cartel,k} P_{-k} \tag{14}$$

and verify it. Given the guess, repeatedly substituting (14), we have

$$P_{+k} = a_{cartel}\left(1 + \sum_{l=1}^{k} b_{cartel,l} \sum_{m=1}^{l} \left(b_{cartel,1}\right)^{m-1}\right) + \sum_{l=0}^{k}\left(\left(\sum_{m=1}^{k-l} \left(b_{cartal,m}\right)^{k-m}\right)\left(\sum_{n=1}^{K-l} b_{cartel,n+l} P_{-n}\right)\right)$$

with $\sum_{m=1}^{0} \left(b_{cartal,m}\right)^{k-m} \equiv 1$. Putting them back into the first-order condition (13) and matching the coefficient gives us the linear solution, and so the guess is verified. ■

MPE Suppose no cartel exists. Then each firm would like to maximize its profit by changing q_i, taking the other firm's output as given. Because the price is determined by

$$P = -\psi \times (x - Q_{fri}) + \psi \times (Q_{-i} + q_i),$$

we can see firm i picks the optimal price, and then produce

$$q_i = x - Q_{fri} + \frac{P}{\psi} - Q_{-i},$$

rationally expecting the other players and the fringe produce Q_{-i} and Q_{fri}, respectively.

Lemma 2 *There exists a linear MPE: each firm i produces $q_i\left(P_{-1:-K}\right) = a^i_{MPE} + \sum_{k=1}^{K} b_{MPE,k} P_{-k} + \gamma^i_{MPE} c_i$.*

Note that the reaction to the price, $b_{MPE,k}$, does not depend on the index of the player.

Proof. Suppose the other players follow an MPE, and so Q_{-i} is determined by $P_{-1:-K}$. Then the optimal cartel supply q_i given Q_{-i} is to maximize

$$V^{i,MPE}\left(P_{-1:-K}\right) = \max_{P} \left(P - c_i\right)\left[x - Q_{fri} + \frac{P}{\psi} - Q_{-i}\right] + \beta V^{i,MPE}\left(P, P_{-1:-K+1}\right).$$

We guess $q_i\left(P_{-1:-K}\right) = a^i_{MPE} + \sum_{k=1}^{K} b_{MPE,k} P_{-k} + \gamma^i_{MPE} c_i$ for each i, and verify this strategy satisfies this value function.

The first-order condition is

$$x - Q_{fri} - Q_{-i} + \frac{2P - c_i}{\psi} + \beta V^{i,MPE}_1\left(P, P_{-1:-K+1}\right) = 0.$$

By the envelope theorem, we have

$$
\begin{aligned}
V_k^{i,MPE}\left(P_{-1:-K}\right) &= \left(P-c_i\right)\left(-b_{fri,k}-\frac{dQ_{-i}}{dP_{-k}}\right)+\beta V_{k+1}^{i,MPE}\left(P,P_{-1:-K+1}\right)\\
&= \left(P-c_i\right)\left(-b_{fri,k}-\left(N-1\right)b_{MPE,k}\right)+\beta V_{k+1}^{i,MPE}\left(P,P_{-1:-K+1}\right)
\end{aligned}
$$

and so

$$
\beta V_1^{i,MPE}\left(P,P_{-1:-K+1}\right)=-\sum_{k=1}^{K}\beta^k\left(P_{+k}-c_i\right)\left(b_{fri,k}+\left(N-1\right)b_{MPE,k}\right). \tag{15}
$$

Hence the first-order condition is equivalent to

$$
P=\frac{|\psi|\left(x-Q_{fri}-Q_{-i}\right)+c_i-|\psi|\sum_{k=1}^{K}\beta^k\left(P_{+k}-c_i\right)\left(b_{fri,k}+\left(N-1\right)b_{MPE,k}\right)}{2}.
$$

Adding them up with respect to i, we have

$$
\begin{aligned}
P &= \frac{|\psi|\left(x-Q_{fri}-\frac{N-1}{N}Q\right)+\bar{c}}{2}\\
&\quad -\frac{|\psi|}{2}\sum_{i\in I}\sum_{k=1}^{K}\beta^k\left(P_{+k}-\bar{c}\right)\left(b_{fri,k}+\left(N-1\right)b_{MPE,k}\right),
\end{aligned} \tag{16}
$$

where $\bar{c}=\frac{1}{N}\sum_i c_i$ is the average marginal cost.

Substituting (16) into $Q=x-Q_{fri}-\frac{P}{|\psi|}$, we have

$$
Q=\frac{N}{N+1}\left(x-Q_{fri}\right)-\frac{\sum_i c_i}{|\psi|\left(N+1\right)}+\frac{\sum_{n=1}^{k}\beta^n\left(NP_{+n}-\sum_i c_i\right)\left(b_{fri,n}+\left(N-1\right)b_{MPE,k}\right)}{N+1} \tag{17}
$$

and

$$
P=|\psi|\left(x-Q_{fri}-Q\right).
$$

The rest of the proof (matching the coefficient) is the same as Lemma 1, and so is omitted.

■

Equilibrium Concept

We consider the following equilibrium based on the trigger strategies (now the punishment to trigger is the linear MPE specified in Lemma 2). Suppose the firms interact without cartels for a while, and they reach the stationary quantity and price of the Markov perfect equilibrium $\bar{P}$.

Cartel Profit Given the state $P_{t-k} = \bar{P}$ for each k, we calculate the sequence of cartel quantities and prices, $Q_t = \bar{Q}\left(\bar{P}, ..., \bar{P}\right)$, $P_t = P\left(\bar{P}, ..., \bar{P}\right)$, $Q_{t+1} = \bar{Q}\left(\bar{P}, ..., \bar{P}, P_t\right)$, $\bar{P}_{t+1} = P\left(\bar{P}, ..., \bar{P}, P_t\right)$, $Q_{t+2} = \bar{Q}\left(\bar{P}, ..., \bar{P}, P_t, P_{t+1}\right)$, $\bar{P}_{t+2} = P\left(\bar{P}, ..., \bar{P}, P_t, P_{t+1}\right)$, and so on. Suppose firm i obtains the market share α_i. Then its cartel profit is $(P_t - c_i)\,\alpha_i Q_t$, $(P_{t+1} - c_i)\,\alpha_i Q_{t+1}$, and so on.

Optimal Non-compliance If the cartel agreement is determined as above, the optimal non-compliance profit in period t' is calculated as follows: given on-path price $\{P_{t'-k}\}_{k=1}^{K}$, the other firms will produce the quantity according to $(1-\alpha_i)\,\bar{Q}\left(\{P_{\tilde{t}-k}\}_{k=1}^{K}\right)$ for the next L periods $\tilde{t} = t', ..., t'+L-1$, and then will switch to the Markov perfect equilibrium $q_i\left(\{P_{\tilde{t}-k}\}_{k=1}^{K}\right)$ from $\tilde{t} \geq t'+L$.

In particular, for each past price $(P_{-1:-k})$ (which determines fringe output Q_{fri} and cartel output Q_{cartel}), we can derive the optimal non-compliance by backward induction.

Suppose firm i deviated $L-1$ periods ago, and so this period is the last one before switching to MPE. Given that firm i has α_i share in the cartel agreement, we have

$$P = -\psi\left(x - Q_{fri}\right) + \psi\left(\alpha_{-i}Q_{cartel} + q_i\right)$$

with $\alpha_{-i} = 1 - \alpha_i$. In other words,

$$q_i = \left(x - Q_{fri}\right) + \frac{P}{\psi} - \alpha_{-i}Q_{cartel}.$$

Because Q_{cartel} is known, selecting q_i is equivalent to selecting P.

The optimal non-compliance given Q_{cartel} is to maximize

$$V^{i,L}\left(P_{-1:-K}\right) = \max_{P}\left(P - c_i\right)\left[x - Q_{fri} + \frac{P}{\psi} - \alpha_{-i}Q_{cartel}\right] + \beta V^{i,MPE}\left(P, P_{-1:-K+1}\right).$$

From (15), we know $V_1^{i,MPE}(P, P_{-1:-K+1})$. Hence the first order condition becomes

$$P = \frac{|\psi|(x - Q_{fri} - \alpha_{-i}Q_{cartel}) + c_i}{2} - \frac{|\psi| \sum_{k=1}^{K} \beta^n (P_{+k} - c_i)(b_{fri,k} + (N-1) b_{MPE,k})}{2}. \tag{18}$$

Because we know the law of motion of $\{P_{+K:+1}\}$ from Lemma 2, we can solve P as a function of Q_{fri}, that is, a function of $P_{-1:-K}$. Putting this solution to the value function, we derive $V^{i,L}(P_{-1:-K})$.

Given $V^{i,L}$, if firm i deviated $L-2$ periods ago, firm i wants to maximize

$$V^{i,L-1}(P_{-1:-K}) = \max_P (P - c_i)\left[(x - Q_{fri}) + \frac{P}{\psi} - \alpha_{-i}Q_{cartel}\right] + \beta V^{i,L}(P, P_{-1:-K+1}).$$

Because we now know $V^{i,L}$, we can solve for the optimal P as a function of Q_{fri}, Q_{cartel}, and $P_{-1:-K}$. Because the former two are functions of $P_{-1:-K}$, we can solve for the optimal P as a function of $P_{-1:-K}$. By backward induction, we solve for optimal non-compliance prices and quantities.

Equilibrium Condition to Start the Cartel The cartel starts from the steady-state price $\bar{P}$ if and only if, for the on-path sequence of cartel quantities and prices, $Q_t = \bar{Q}(\bar{P}, ..., \bar{P})$, $P_t = P(\bar{P}, ..., \bar{P})$, $Q_{t+1} = \bar{Q}(\bar{P}, ..., \bar{P}, P_t)$, $\bar{P}_{t+1} = P(\bar{P}, ..., \bar{P}, P_t)$, $Q_{t+2} = \bar{Q}(\bar{P}, ..., \bar{P}, P_t, P_{t+1})$, $\bar{P}_{t+2} = P(\bar{P}, ..., \bar{P}, P_t, P_{t+1})$, there are no i and τ such that

$$\underbrace{V(P_{\tau-1:\tau-K})}_{\text{the discounted sum of cartel profits}} < \underbrace{V^{i,1}(P_{\tau-1:\tau-K})}_{\text{the discounted sum of profits when firm } i \text{ deviates in period } \tau} \tag{19}$$

Empirical Analysis

We specify $K = 2$ and regress $Q_{fri,t}$ on P_{t-2} and P_{t-1} to estimate the fringe supply function. Given this estimate, we see if there exists any share profile $(\alpha_i)_{i \in I}$ with which (19) holds and verify no feasible $(\alpha_i)_{i \in I}$ exists for each $\beta \leq .9$.

Compared to the myopic optimal quantity (not taking into account the fringe's reaction), the optimal price will be lower because the higher price will enhance the future entry of the Chinese firms. This lower optimal price makes the continuation payoff under the collusion small, which makes collusion incentive incompatible.

C.2 Renegotiating Quotas (continued from section 6.2)

We fix the share of each firm at the level of pre-cartel periods in the baseline analysis. One may wonder if renegotiating and coming up with new shares with which the incentive compatibility constraint is satisfied was possible.

As mentioned in Appendix C.1, coming up with a share to satisfy the incentive compatibility of cartels is impossible.[52] In particular, for each $\beta \leq .9$, to maintain Roche's incentive, they have to give share $\alpha_i > .8$ to Roche (i.e., with $\alpha_i = .8$, Roche's incentive compatibility is violated). At the same time, to maintain Takeda's incentive, they have to give $\alpha_i > .2$ to Takeda (i.e., with $\alpha_i = .2$, Takeda's incentive compatibility is violated). These are incompatible. Hence, we conclude the renegotiation of shares does not allow the firms to reform the cartel.

C.3 Cartel Price below Monopoly Level (continued from section 6.3)

Section 5.3 showed the cartel pricing in the data converged to theoretical monopoly levels in all four vitamin markets within a year or two. Hence, our subsequent analysis treated cartel prices and monopoly prices interchangeably. However, existing studies suggest real-world cartels do not always entail monopoly prices. This section investigates the sensitivity of our incentive estimates with respect to the level of cartel pricing. We find alternative prices below the monopoly level could not have relaxed the IC constraint in any significant way.

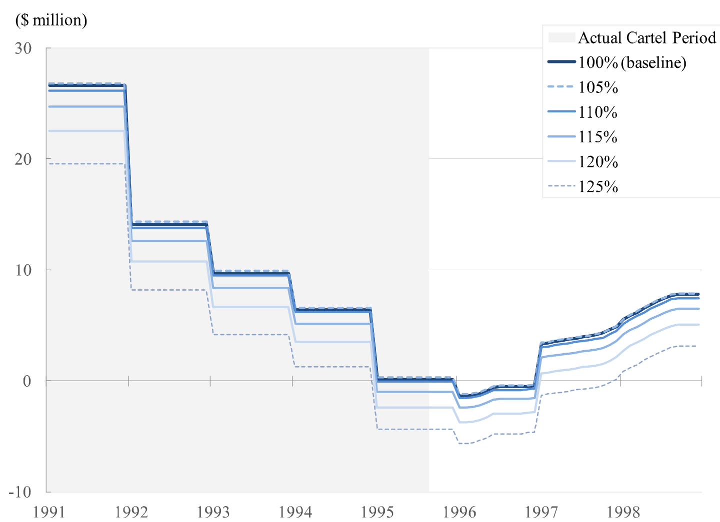
The cartel's collective outputs under static Nash surpasses its perfectly coordinated (i.e., monopoly) level by 28% $\sim$ 36% throughout the sample period. Instead of assuming perfect coordination, suppose the cartel's collective outputs exceeded the monopoly level by 5% $\sim$ 25%.

Figure 15 shows the IC constraint would generally become more binding if the cartel coordinated to achieve less-than-monopoly levels. The bold line (100%) is the baseline estimate with $\beta = 0.7$, whereas the five other reflect greater output levels of the cartel (105% $\sim$ 125%), with the labels indicating the percentages of the cartel's monopoly output. The 105% version performs slightly better than the baseline, but the improvement is barely visible. All other versions underperform the baseline. Thus, the cartel did not have a reason to set a target price below the monopoly level.

[52] Here, we do the exercise of varying shares in the specification of Appendix C.1 (endogenous fringe supplies). We can do the same exercise with the static belief about the fringe supply, simply varying the market share in the baseline analysis.

These outcomes are reasonable for the following reasons. On the one hand, $\frac{dV^C}{dP} = 0$ in the neighborhood of the monopoly price because $\frac{d\pi^C}{dP} = 0$ by construction and $\frac{d\pi^N}{dP} = 0$ regardless of the cartel price. On the other hand, the sign of $\frac{dV^D}{dP}$ is not obvious, and hence ΔV might increase with a small change from the monopoly price (e.g., when the cartel produces 105% of the monopoly output). However, when the price departs significantly from the monopoly level (e.g., when the cartel produces 110% $\sim$ 125% of the monopoly output), its first-order (negative) effect on π^C kicks in to drive down V^C along with ΔV.

Figure 15: IC Constraint with Cartel Prices below Monopoly Level



Note: The legend labels each line by the cartel's *output* as a percentage of the monopoly level. See text.

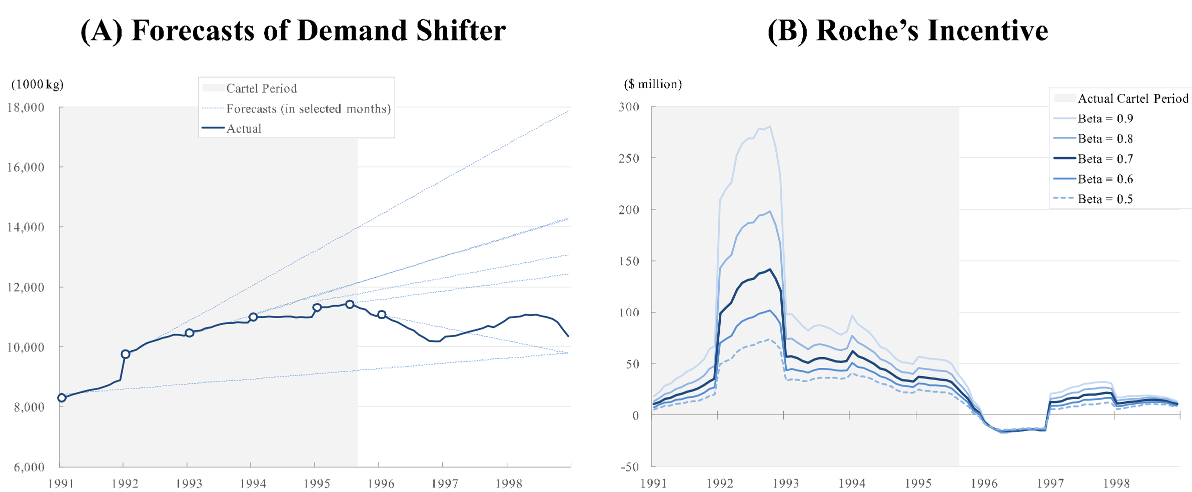
C.4 Adaptive Expectations on Future Demand (continued from section 6.4)

In section 5.4, we showed the estimates of Roche's incentive to collude under the assumption of rational expectations. We conducted a series of sensitivity analyses with multiple different specifications in the spirit of adaptive expectations but obtained quantitatively similar outcomes on the predicted timing of the cartel's breakup.

Figure 16 shows additional results under the adaptive expectations setting with a one-year window (i.e., $K = 13$), which makes the firms' demand forecast highly sensitive to their contemporaneous observations. Our model's prediction moves closer to the historical timing of the collapse in August 1995. Moreover, the prediction becomes less sensitive to the calibrated level of discount factor, β. Thus, the alternative assumption of adaptive expectations would make our model's prediction more accurate in terms of timing and more robust with respect to β.

The reason is that the demand $\left(\tilde{X}_t\right)$ stops growing and starts declining in the middle of 1995. Firms with rational expectations would correctly foresee the temporary nature of such a downturn and keep colluding for a while, but adaptive expectations (with reasonably short memories) would place higher weights on such declining trends and make the firms more pessimistic about the net gains from collusion, which breaks the incentive compatibility condition exactly at the time when the cartel collapsed in the data. This mechanism is the one through which adaptive expectations make our results stronger and more precise. Despite such favorable results under adaptive expectations, however, we still prefer rational expectations as our baseline specification on the factual grounds. Historical evidence suggests the firms expected a secular growth trend with at least five to 10 years of time horizons.

Figure 16: Roche's Incentive under Adaptive Expectations



C.5 Log Linear Demand (continued from section 6.5)

Consider the following log-linear demand:

$$\log Q_t^D = \alpha_0' + \alpha_1' \log P_t + \alpha_2' \log X_t + \varepsilon_t', \tag{20}$$

where $\alpha_1' = \frac{\partial \log Q}{\partial \log P}$ can be interpreted as the price-elasticity of demand. Regardless of the functional form of the demand, the Cournot FOC remains unchanged and identifies $\frac{\partial P}{\partial Q}$.

Thus, the estimates of the slope and the elasticity of the demand does not depend on the functional form. By contrast, the effective demand shifter is recovered as the residual of the demand function,

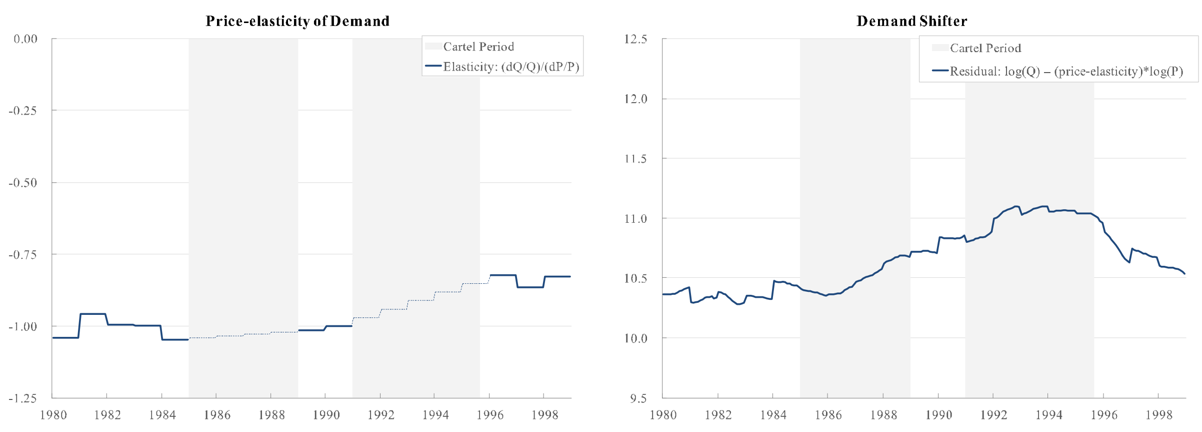
$$\begin{aligned} \tilde{X}'_t &\equiv \alpha'_0 + \alpha'_2 \log X_t + \varepsilon'_t \\ &= \log Q_t - \hat{\alpha}'_1 \log P_t, \end{aligned} \tag{21}$$

and hence it depends on the specification of the latter by definition.

Figure 17 (left) shows the price elasticity of demand. Outside the cartel period, its estimates are identical to the baseline results. During the cartel period, we interpolated $\alpha_1 = \frac{\partial Q}{\partial P}$ in our baseline analysis, whereas we interpolate $\alpha'_1 = \frac{\partial \log Q}{\partial \log P}$ in the current analysis because the two specifications slightly differ from each other in terms of what constitutes the main parameter (i.e., α_1 or α'_1).

Figure 17 (right) shows the effective demand shifter under the log-linear specification is largely devoid of a growth trend. Its level in the late 1990s is almost as low as that in the early 1980s. This feature is at odds with the testimony of industry experts in both the court documents and the government reports, which unanimously characterized vitamin demand (including vitamin C) as growing steadily. We suspect the apparent lack of demand growth is simply an artifact of the log-linear functional form, and prefer our baseline, linear specification.

Figure 17: Log Linear Demand Estimates



C.6 Long-term Contracts (continued from section 6.6)

Suppose the renewal of 12-month contracts has the following structure. The demand function in period t,

$$Q_t^D = \alpha_0 + \alpha_1 P_t + \alpha_2 X_t + \varepsilon_t,$$

means that, given price P_t, total demand Q_t^D is to be delivered over the next 12 periods, (i.e., $\frac{Q_t^D}{12}$ is delivered to the consumer from period t to period $t+11$). That is, the quantity $q_{i,t}$ "sold" by firm i in period t will be delivered over the next 12 periods.[53] The per-unit payment is fixed at P_t in period t, and the payment is made upon delivery. Hence, the discounted profit from $q_{i,t}$ with P_t is

$$\frac{1}{12}P_t q_{i,t} + \frac{1}{12}\beta P_t q_{i,t} + \cdots + \frac{1}{12}\beta^{11} P_t q_{i,t} = \frac{1-\beta^{12}}{1-\beta}\frac{P_t q_{i,t}}{12}.$$

Because $\frac{1-\beta^{12}}{1-\beta}\frac{1}{12}$ is a constant, the optimal cartel quantity, the optimal deviation quantity, and the static Nash quantity are the same. Hence, firm i's payoff under compliance is

$$\begin{aligned}
&\frac{1-\beta^{12}}{1-\beta}\pi^C_{i,\tau|t} + \sum_{k=1}^{\infty}\beta^{12k}\frac{1-\beta^{12}}{1-\beta}\pi^C_{i,\tau+12k|t} \\
&+\beta\left(\frac{1-\beta^{12}}{1-\beta}\pi^C_{i,\tau+1|t} + \sum_{k=1}^{\infty}\beta^{12k+1}\frac{1-\beta^{12}}{1-\beta}\pi^C_{i,\tau+1+12k|t}\right) \\
&+\cdots+\beta^{11}\left(\frac{1-\beta^{12}}{1-\beta}\pi^C_{i,\tau+11|t} + \sum_{k=1}^{\infty}\beta^{12k+11}\pi^C_{i,\tau+11+12k|t}\right),
\end{aligned}$$

[53] Note the demand is properly scaled. Suppose a firm i has $q_{i,t} = q$ for all t. Suppose $q_{i,t}$ is delivered over periods t-$t+11$, $\frac{q_{i,t}}{12}$ for each month. Then $q_{i,1} = q$ is delivered over periods 1-12, $q_{i,2} = q$ is delivered over periods 2-13, and so on. Hence, the total quantity delivered is q per month in the steady state.

while firm i's payoff under non-compliance is

$$\frac{1-\beta^{12}}{1-\beta}\pi^{D}_{i,\tau|t}+\sum_{k=1}^{\infty}\beta^{12k}\frac{1-\beta^{12}}{1-\beta}\pi^{P}_{i,12k+1|t}$$
$$+\beta\left(\frac{1-\beta^{12}}{1-\beta}\pi^{D}_{i,\tau+1|t}+\sum_{k=1}^{\infty}\beta^{12k+1}\frac{1-\beta^{12}}{1-\beta}\pi^{P}_{i,\tau+1+12k|t}\right)$$
$$+\beta^{2}\left(\frac{1-\beta^{12}}{1-\beta}\pi^{D}_{i,\tau+2|t}+\sum_{k=1}^{\infty}\beta^{12k+2}\frac{1-\beta^{12}}{1-\beta}\pi^{P}_{i,\tau+2+12k|t}\right)$$
$$+\beta^{3}\left(\frac{1-\beta^{12}}{1-\beta}\pi^{P}_{i,\tau+4|t}+\sum_{k=1}^{\infty}\beta^{12k+1}\frac{1-\beta^{12}}{1-\beta}\pi^{P}_{i,\tau+1+12k|t}\right)$$
$$+\cdots+\beta^{11}\left(\frac{1-\beta^{12}}{1-\beta}\pi^{P}_{i,\tau+11|t}+\sum_{k=1}^{\infty}\beta^{12k+11}\frac{1-\beta^{12}}{1-\beta}\pi^{P}_{i,\tau+11+12k|t}\right).$$

Rearranging, the equilibrium condition is

$$\sum_{k=1}^{\infty}\beta^{k-1}\pi^{C}_{i,\tau+k-1|t}\geq\sum_{k=1}^{3}\beta^{k-1}\pi^{D}_{i,\tau+k-1|t}+\sum_{k=4}^{\infty}\beta^{k-1}\pi^{P}_{i,\tau+k-1|t}$$

for each i and $\tau\geq t$. That is, $\min_{i,\tau\geq t}\left(V^{C}_{i,\tau|t}-V^{D}_{i,\tau|t}\right)\geq 0$, as in (4).

Appendix D: Merger Simulation

D.1 Merger May Increase or Decrease Collusive Incentives

The cartel is sustainable if and only if the deviation gain is less than the continuation payoff loss of deviations. For simplicity, let us assume that $P = a - bQ$ and each firm has the same constant marginal cost c. Let N be the number of firms. Moreover, the cartel agreement is for each firm to produce $\frac{a-c}{2bN}$ (a quantity to maximize the joint profit, equally split between the cartel members), and the punishment is a permanent reversion to the static Nash equilibrium.

A simple calculation derives

1. The on-path payoff is $\frac{(a-c)^2}{4bN}$.
2. The punishment payoff is $\frac{(a-c)^2}{(N+1)^2 b}$.
3. The deviation gain is $\left(\frac{N+1}{N}\right)^2 \frac{(a-c)^2}{16b}$.

All of them are decreasing in N, but the on-path payoff (determining $V^1_{i,\tau|t}$ and $V^4_{i,\tau|t}$) is linear in $\frac{1}{N}$; the punishment payoff (corresponding to $V^2_{i,\tau|t}$) is linear in $\frac{1}{(N+1)^2}$; and the deviation gain is linear in $\left(\frac{N+1}{N}\right)^2$. For example, with $a = 9, b = .1, \beta = .9$ (annual discount factor), the left-hand side minus the right-hand side of (9),

$$\left(1+\beta+\beta^2\right)\left(\left(\frac{N+1}{N}\right)^2 \frac{(a-c)^2}{16b} - \frac{(a-c)^2}{4bN}\right) - \sum_{\tau=3}^{\infty} \beta^{\tau} \left(\frac{(a-c)^2}{4bN} - \frac{(a-c)^2}{(N+1)^2 b}\right),$$

is not monotone in N.

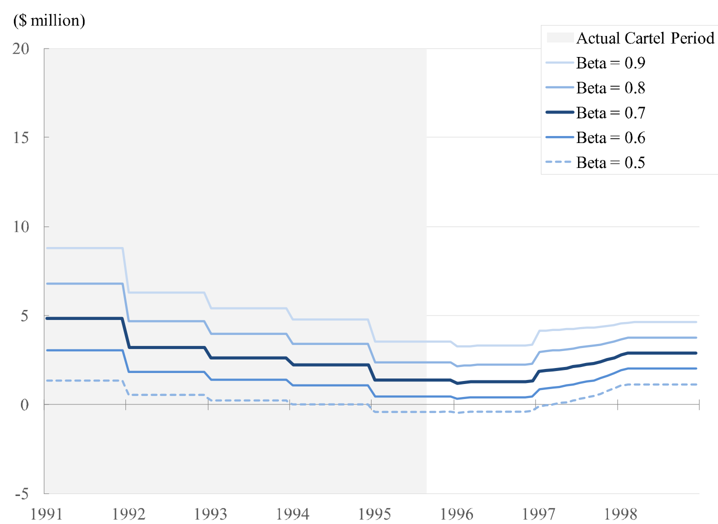
D.2 Roche's Quota Concession

As sections 3.2 and 5.4 described, Roche ceded some of its "historically entitled" quota to Takeda, BASF, and E. Merck at the beginning of the vitamin C cartel. This "sweetening" of the deal is no longer applicable in our merger counterfactual in section 7.2, because the main reason for the original sweetening was to defuse the tension between Takeda and BASF, the merging parties. Once BASF takes over Takeda's vitamin business, a single officer will represent the merged entity (i.e., BASF's managers running former Takeda plants) at the cartel meeting. BASF was the closest partner of Roche and the two leading colluders rarely fought each other (EC 2003). Thus, the BASF-Takeda merger obviates the need for Roche to play the role of a mediator and give up its quota.

Our main finding from the merger simulation is robust to alternative specifications of the counterfactual quotas. Suppose the post-merger integration of former Takeda plants encountered a major difficulty and Takeda's aggressive Osaka culture permeated the new management under BASF. Under such a pessimistic scenario, Roche would still need to concede what it gave to BASF (and E. Merck) in reality.[54]

Figure 18 shows the BASF-Takeda merger before 1991 would have saved the vitamin C cartel even if we incorporate Roche's historical concession to BASF and E. Merck.[55] The $\beta = 0.5$ case is an exception in that it predicts the violation of the IC constraint in 1994, but would not reverse our overall finding.

Figure 18: Roche's IC Constraint If It Concedes Quota



D.3 Aggressive Fringe Supply

Our simulation of the merger's coordinated effect in section 7.2 assumed the Chinese fringe output was the same as its actual level in each year. This section considers alternative

[54] If we make Roche give up as much quota as it did in reality (i.e., including what it gave to Takeda), Roche's on-path continuation value ($V^1_{\tau|t}$) would become the same as that in the baseline model and the merger counterfactual would become mostly meaningless.

[55] We calculate Roche's historical concession to BASF and E. Merck was 3.12% points and 0.15% points of the intra-cartel market shares, respectively, based on the difference between the average intra-cartel market shares between 1991 and 1994 (i.e., the cartel period) and 1989 (i.e., the pre-cartel period). We do not use the 1990 data, because the cartel negotiation had already begun in the vitamin C market.

assumptions about China and assesses the sensitivity of our main results.

Figure 14 in Appendix C.1 ("Endogenous Fringe") shows two upward-sloping supply curves, plotting the total fringe output against the lagged price (left) and the concurrent price (right), respectively. The fringe output strongly correlates with the lagged price but only modestly with the current price, which is consistent with the institutional background of the Chinese suppliers' staggered entry and expansion in section 3.3.

Regardless of the choice between the lagged and the current prices, however, these fringe supply curves do not seem to capture the actual output level in 1998, which is much higher than (i.e., to the right of) the fitted regression lines. Our merger counterfactual in Table 3 entails the average 1998 prices in the range of approximately $9 $\sim$ $12, which should translate into the fringe output of less than $10,000$MT/year according to these supply curves. The actual 1998 fringe output was more than double that level, exhibiting significant downward-rigidity of output adjustments among the Chinese SOEs.

Consequently, we have chosen to simulate an extreme version of such downward rigidity by assuming the fringe output in 1998 is fixed at its peak level in 1996. This pessimistic scenario is a highly ad hoc specification but a useful one for the sake of robustness check, because the estimated fringe supply curves predict unrealistically low outputs and would inflate our estimates of the coordinated effect. Under this worst-case scenario (from the cartel's perspective), the average 1998 price would be $11.28 instead of $11.58 in our main result. The change from the actual (i.e., no merger) level is 15.1% instead of 18.1%. Thus, our conservative assumption about China in this section would make the coordinated effect of merger less pronounced, but the overall impact is still an order of magnitude higher than that of the unilateral effect alone (1.7%).

References

[1] Abreu, D. 1986. "Extremal Equilibria of Oligopolistic Supergames," *Journal of Economic Theory*, 39, 191-225.

[2] Abreu, D., D. Pearce, and E. Stacchetti. 1990. "Toward a Theory of Discounted Repeated Games with Imperfect Monitoring," *Econometrica*, 58, 1041-1063.

[3] Asker, John. 2010. "A Study of the Internal Organization of a Bidding Cartel." *American Economic Review*, 100 (3): 724–762.

[4] Awaya, Y. and Vijay Krishna. 2016. "On Communication and Collusion." *American Economic Review*, 106(2): 285-315.

[5] Bernheim, B. Douglas. 2002a. "Expert Report of B. Douglas Bernheim," *In Re: Vitamins Antitrust Litigation*, MDL No. 1285, Misc 99-0197.

[6] Bernheim, B. Douglas. 2002b. "Rebuttal Expert Report of B. Douglas Bernheim," *In Re: Vitamins Antitrust Litigation*, MDL No. 1285, Misc 99-0197.

[7] Bernheim, B. Douglas. 2008. "Expert Report of B. Douglas Bernheim, Ph.D.," *In Re: Vitamin C Antitrust Litigation*, MDL No. 1738.

[8] Chen, Yuyu, Mitsuru Igami, Masayuki Sawada, and Mo Xiao. 2017. "Privatization and Productivity in China." Manuscript, Yale University.

[9] Connor, John M. 2007. *Global Price Fixing*, second edition. Heidelberg: Springer.

[10] Connor, John M. 2008. "The Great Global Vitamins Conspiracies, 1985-1999." SSRN working paper.

[11] de Roos, Nicolas. 2001. "Collusion with a Competitive Fringe: An Application to Vitamin C." Manuscript, Yale University.

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

[13] Ellison, Glenn. 1994. "Theories of cartel stability and the Joint Executive Committee." *RAND Journal of Economics*, 25 (1): 37–57.

[14] European Commission. 2003. "Case COMP/E-1/37.512 – Vitamins," *Official Journal of the European Communities*.

[15] Fudenberg, D., Y. Ishii, and S. D. Kominers. 2014. "Delayed-Response Strategies in Repeated Games with Observation Lags," *Journal of Economic Theory*, 150, 487-514.

[16] Fudenberg, D., D. Levine, and E. Maskin. 1994. "The Folk Theorem with Imperfect Public Information," *Econometrica*, 62, 997-1039.

[17] Fudenberg, D. and E. Maskin. 1986. "The Folk Theorem in Repeated Games with Discounting or with Incomplete Information," *Econometrica*, 533-554.

[18] Genesove, D. and W. P. Mullin. 1998. "Testing Static Oligopoly Models: Conduct and Cost in the Sugar Industry, 1890-1914," *RAND Journal of Economics*, 29 (2): 355–377.

[19] Green, Edward J., and Robert H. Porter. 1984. "Noncooperative Collusion under Imperfect Price Information." *Econometrica*, 52 (1): 87–100.

[20] Greif, A. 2006. *Institutions and the Path to the Modern Economy*, New York: Cambridge University Press.

[21] Harrington, Joseph E. 2006. "How Do Cartels Operate?" *Foundations and Trends in Microeconomics*, 2 (1): 1–105.

[22] Harrington, J.E. and A. Skrzypacz. 2007. "Collusion under Monitoring of Sales," *RAND Journal of Economics*, 38, 314-31.

[23] Harrington, Joseph E., and Andrzej Skrzypacz. 2011. "Private Monitoring and Communication in Cartels: Explaining Recent Collusive Practices." *American Economic Review*, 101: 1–25.

[24] Kandori, M. 2002. "Introduction to Repeated Games with Private Monitoring," *Journal of Economic Theory*, 102, 1-15.

[25] Kandori, M., and Shinya Obayashi. 2014. "Labor union members play an OLG repeated game." *Proceedings of the National Academy of Sciences of the United States of America*, 111 (3): 10802–10809.

[26] Kaplow, Louis. 2013. *Competition Policy and Price Fixing*, Princeton Univ. Press.

[27] Kawai, Kei, and Jun Nakabayashi. 2015. "Detecting Large-Scale Collusion in Procurement Auctions." Manuscript, UC Berkeley.

[28] Kobayashi, Hajime, and Katsunori Ohta. 2012. "Optimal collusion under imperfect monitoring in multimarket contact." *Games and Economic Behavior*, 76 (2): 636-647.

[29] Kovacic, William E., Robert C. Marshall, Leslie M. Marx, and Steven P. Schulenberg. 2009. "Quantitative Analysis of Coordinated Effects." *Antitrust Law Journal*, 76: 397–430.

[30] Kreps, D., P. Milgrom, J. Roberts, and R. Wilson. 1982. "Rational Cooperation in the Finitely Dilemma Repeated Prisoners' Dilemma," *Journal of Economic Theory*, 27, 245–252.

[31] Kreps, D.M., and J.A. Scheinkman. 1983. "Quantity precommitment and Bertrand competition yield Cournot outcomes." *The Bell Journal of Economics*, 326–337.

[32] Kreps, D. and R. Wilson. 1982. "Reputation and Imperfect Information," *Journal of Economic Theory*, 27, 253–279.

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

[34] Levenstein, Margaret, and Valerie Suslow. 2011. "Breaking Up Is Hard to Do: Determinants of Cartel Duration." *Journal of Law and Economic*s, 54 (2): 455–492.

[35] Levenstein, Margaret, and Valerie Suslow. 2014. "Cartels and Collusion: Empirical Evidence." In Roger D. Blair and D. Daniel Sokol (eds.), *The Oxford Handbook of International Antitrust Economics*, vol. 2.

[36] Li, Wenjian. 2002. "China's chemical materials industry: Facing the historic opportunity of industrial development," Guotai Junan Securities Research Institute, Research Special Issue No. 124, August 18 [in Chinese].

[37] Mailath, G.J. and L. Samuelson. 2006. *Repeated Games and Reputations*, Oxford: Oxford University Press.

[38] Marshall, Robert C., and Leslie M. Marx. 2014. *The Economics of Collusion: Cartels and Bidding Rings*. Cambridge, Massachusetts: The MIT Press.

[39] Marshall, Robert C., Leslie M. Marx, and Matthew E. Raiff. 2008. "Cartel price announcements: The Vitamins industry." *International Journal of Industrial Organization*, 26: 762–802.

[40] Matsushima, Hitoshi. 2001. "Multimarket Contact, Imperfect Monitoring, and Implicit Collusion." *Journal of Economic Theory*, 98 (1): 158-178.

[41] Milgrom, P. and J. Roberts. 1982. "Predation, Reputation and Entry Deterrence," *Journal of Economic Theory*, 27, 280–312.

[42] Miller, Nathan H., and Matthew C. Weinberg. 2015. "Mergers Facilitate Tacit Collusion: Empirical Evidence from the U.S. Brewing Industry." Manuscript, Georgetown University.

[43] Porter, Robert H. 1983. "A study of cartel stability: the Joint Executive Committee, 1880–1886." *Bell Journal of Economics*, 301–314.

[44] Porter, Robert H., and J. Douglas Zona. 2008. "Collusion." *Antitrust Law Journal*, 2: 1069–1084.

[45] Rotemberg, J.J. and G. Saloner. 1986. "A Supergame-Theoretic Model of Price Wars During Booms," *American Economic Review*, 76, 390-407.

[46] Sekiguchi, Tadashi. 2015. "Multimarket Contact under Demand Fluctuations: A Limit Result." *mimeo.*

[47] Stigler, George J. 1964. "A Theory of Oligopoly." *Journal of Political Economy*, 72 (1): 44–61.

[48] Takeda Pharmaceutical Industry. 1983. *Takeda Nihyakunen Shi* (200 years of company history), Osaka: Takeda Pharmaceutical Industry [in Japanese].

[49] Takeda Pharmaceutical Industry, *Yuka shoken hokokusho* (regulatory filings of financial and operational reports), 1980-2005 [in Japanese].

[50] The United Kingdom Competition Commission. 2001. *BASF and Takeda Chemical Industries Ltd: A Report on the Acquisition by BASF AG of Certain Assets of Takeda Chemical Industries Ltd.* London: UK Competition Commission.

[51] Xiang, Anbo, Zhongjun Zhang, Xiaohong Chen, and Mariko Watanabe. 2007. "Chuugoku iyakuhin sangyo: Sangyo no zentaizou," in Kensuke Kubo (ed.), *Nihon no generic iyakuhin shijou to indo chuugoku no seiyaku sangyo*, IDE-JETRO [in Japanese].

[52] Yamaguchi, Hidemaru. 2000. "Takeda Pharmaceutical Company (4502): Revised earning forecast and target price." Nikko Salomon Smith Barney, July 31, 2000.