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

Mitsuru Igami†  Takuo Sugaya‡

May 21, 2019

Abstract

Do mergers help or hinder collusion? This paper studies the stability of the vitamin cartels in the 1990s and presents a repeated-games approach to quantify “coordinated effects” of a merger. We use data and direct evidence from American courts and European agencies to show the incentive compatibility constraint (ICC) of the short-lived vitamin C cartel was likely to be violated when it actually collapsed in 1995, whereas the ICCs of the long-lived cartels (vitamins A and E, and beta carotene) were satisfied until the prosecution in 1999. Simulations suggest some mergers could have prolonged the vitamin C cartel, but others could have further destabilized it, because both the direction and magnitude of “coordinated effects” depend not only on the number of firms but also on their cost asymmetry.

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

JEL classifications: D43, L13, L41.

---


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

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

One of the most fundamental ideas of repeated-game theory is that cooperation is sustainable if and only if it is incentive compatible. Theory provides a “checklist” of relevant factors that affect the sustainability of collusion, including the number of competitors, the degree of symmetry among them, demand growth, and fringe supply. However, predicting the likelihood of collusion on this basis alone would be difficult, because a given market will have some characteristics that facilitate collusion and others that hinder collusion.\(^1\) In particular, a merger poses a practical challenge because it affects many of these factors at the same time (e.g., reducing the number of firms may facilitate collusion, but a merger can also make the remaining firms more asymmetric, which would hinder collusion), which makes its impact theoretically ambiguous.\(^2\) For these reasons, both academics and practitioners have recognized the need for a quantitative analysis of “coordinated effects” (i.e., the effects of mergers on the possibility and durability of collusion) based on a structural model and prospective merger simulations.\(^3\)

This paper presents a structural analysis of coordinated effects with a concrete example of the vitamin cartels in the 1990s, which are one of the largest antitrust cases in history.\(^4\) We quantify the cartels’ incentive compatibility constraints (ICCs) in a model of collusion by using data and direct evidence from American courts and European agencies. The measurement of the ICCs allows us to assess the empirical relevance of the theory of collusion more directly than in the existing research. Moreover, the quantitative ICCs help us disentangle various factors that affect the sustainability of collusion, and address an important policy question: Will a merger create a situation in which collusion is significantly easier to sustain? Our results suggest the combination of a simple model and “typical” data that are available to antitrust agencies can both (i) explain what happened to the vitamin cartels and (ii) provide an analytical tool to assess coordinated effects of mergers.

The vitamins case is like a laboratory of cartel stability. Roche, a Swiss drug company, cooperated with 20 other vitamin makers around the world and cartelized 16 different product categories at the beginning of the 1990s. Some cartels collapsed internally in 1994 or 1995

\(^1\)Ivaldi, Jullien, Rey, Seabright, and Tirole (2003) summarize the theoretical “checklist.” Their report to the European Commission (EC) explains why mergers’ effects on collusion are ambiguous, and advocate for a structural analysis to quantify the effects of these factors. Levenstein and Suslow (2006) and Harrington (2006) summarize stylized facts.

\(^2\)Compte, Jenny, and Rey (2002) and Vasconcelos (2005) study these theoretical relationships.

\(^3\)Merger simulations to assess coordinated effects have been proposed by economists at the UK Competition Commission (Davis and Huse [2010]) and the Italian Competition Authority (Sabbatini [2006]), as well as Kovacic, Marshall, Marx, and Schlenenberg (2009) in the American context.

\(^4\)See https://www.justice.gov/atr/sherman-act-violations-yielding-corporate-fine-10-million-or-more.
(e.g., vitamin C) allegedly because of a slowdown of demand growth and a sudden expansion of fringe supply; others survived until the prosecution in 1999 (e.g., vitamins A and E, and beta carotene). Figure 1 summarizes the price trends in the four largest categories.

Figure 1: Cartels and Vitamin Prices

Note: Roche’s monthly average US transaction prices ($/kg) are rescaled with January 1995 as the base period. The vitamin C cartel collapsed internally in August 1995, whereas the other three cartels continued operations until 1999, when the US government prosecuted them.


Evidence from American courts and European agencies suggests each cartel used a quota agreement based on percentage shares of the global market (on a tonnage basis) in the pre-cartel period. They met every three months to verify each other’s self-reported sales quantity by checking government statistics on imports and exports. Their communication record indicates they shared an understanding that the violation of the quota would terminate the agreement and bring back competitive prices for the foreseeable future. They seemed successful in predicting the steady growth of world demand, but failed to predict a sudden expansion of fringe supply in some categories (e.g., vitamin C).

Based on this documented evidence, we characterize the spot market of each vitamin with a quantity game, and use quantitative information in Bernheim (2002a) to estimate its key parameters, in the four markets for which the data quality is the highest: vitamins A, C, and E, and beta carotene. Following the details in the European Commission’s report (EC 2003), we model the cartel agreement as a quota system to maximize the collective profits
(given the demand curve, fringe supply, and the leading firm’s cost), and assume the firms use trigger strategies with the threat of reversion to static Nash if someone’s non-compliance is confirmed by trade statistics at the quarterly meetings (i.e., “perfect” monitoring with a three-month delay). We conduct sensitivity analysis with alternative model specifications concerning fringe supply, quotas, cartel pricing, monitoring lag, expectations, demand, and long-term contracts.

We present three sets of results. First, the ICC of the short-lived vitamin C cartel was likely to be violated at the time of its actual collapse in 1995, whereas the ICCs of the long-lived cartels for vitamins A and E and beta carotene were substantially less stringent through 1999. Thus, we show evidence in support of the theory of collusion regarding its central prediction that if a cartel collapses internally, it does so because the environment has changed so that the ICC is no longer satisfied. We use this argument and counterfactual simulations to quantify the effects of demand growth and fringe supply on the stability (or collapse) of the vitamin C cartel.

Second, we measure the impact of a merger on cartel stability, by simulating a hypothetical situation in which BASF’s acquisition of Takeda’s vitamin assets (which actually happened in 2001) had been consummated before the beginning of the vitamin C cartel in 1991. We show this merger would have relaxed the ICCs only negligibly. Simulations with varying degrees of synergy suggest the impact of efficiency gains on coordinated effects is highly non-monotonic. Mildly positive synergies, such as a 5% reduction in BASF-Takeda’s marginal cost, would have relaxed the ICCs further. However, larger efficiency gains (e.g., over 10%) would have destabilized the cartel, because synergies make the merged entity’s collusive incentive more aligned with those of low-cost rivals but less aligned with those of high-cost rivals.

Third, although mergers reduce the number of firms and increase market-share concentrations, which are usually associated with less competitive market structures, our simulations suggest mergers (from four to three or two firms) could also simultaneously lead to less collusive incentive structures. Specifically, mergers that eliminate the highest-cost firms make the remaining firms’ cost profiles more symmetric and relax their ICCs, whereas mergers that eliminate medium-cost firms polarize the cost profile and destabilize collusion. Thus, concentration may increase market power in terms of unilateral effects, but both the sign and the magnitude of coordinated effects critically depend on cost asymmetry. These findings highlight the feasibility and desirability of quantitative analysis to inform merger enforcement with respect to coordinated effects.
This paper relates to three literatures. First, the game-theoretic literature on collusion has developed and tested the predictions of repeated-game theory. Many papers focused on the timing of price wars, including Porter (1983), Green and Porter (1984), Rotemberg and Saloner (1986), Bresnahan (1987), Slade (1987, 1992), Haltiwanger and Harrington (1991), Ellison (1994), and Borenstein and Shepard (1996). By contrast, we study cartel stability in the longer run in terms of the ICC and its determinants. We clarify conditions under which collusion is sustainable in a structural empirical model, and present more direct evidence in support of the theoretical prediction by quantifying ICCs.

Second, a large literature exists on the internal operations of cartels. More recently, Asker (2010) studied the bidding ring of stamp dealers in the 1990s; Harrington and Skrzypacz (2011) explain stylized facts by a new theoretical model; and Clark and Houde (2013) studied the retail gasoline cartels in Quebec (2005–2006). Our work is similar to these papers in studying a relatively recent case and in using a structural model, but diverges from them with our focus on the measurement of ICCs and coordinated effects of mergers.

Third, coordinated effects of mergers have been an active area of antitrust policy discussion, but empirical evidence has been scarce, with the notable exception of Miller and Weinberg’s (2017) study of a beer merger case. We complement their retrospective merger analysis by showcasing a prospective merger simulation, which has become common in quantifying “unilateral effects” of mergers (e.g., Werden and Froeb [1994], and Nevo [2000]) but not for coordinated effects. We extend these efforts by using a repeated-games framework. Although the analysis becomes inherently dynamic, our data requirement is similar to that of usual (static) merger analysis and does not impose an additional burden.

2 The Vitamin Cartels, 1990–1999

The vitamins case has been studied extensively, and we refer the reader to excellent books by Connor (2007) and Marshall and Marx (2014) for its general background. This section focuses on more specific features of the cartels that are essential for our modeling purposes. Appendix A provides further details on the institutional background.

---


6 Our work is also categorically related to Fershtman and Pakes (2000), who proposed a Markovian dynamic oligopoly model with collusion, and de Roos (2006), who applied their model to the lysine cartel.
2.1 The Cartels

EC’s (2003) judgment provides detailed information on the internal organization of the vitamin cartels. The EC assesses the quality of the evidence is high.\(^7\)

Before the beginning of collusion, the prices of most vitamins were low. Firms blamed each other’s “price offensive” and “aggressive pricing policy” in vitamins A and E. At the beginning of the vitamin C cartel, Roche explained to BASF that “it had been attempting to raise the price level for the past 4 or 5 years” with little success.\(^8\) In the summer of 1989, the heads of the vitamins divisions of Roche and BASF, as well as Rhône-Poulenc’s (RP) head of the Animal Nutrition division, met in Basel and Zurich to start cartels for vitamins A and E. They agreed to freeze market shares at the 1988 levels for the foreseeable future: “The market for 1990 was estimated and the forecast agreed; the percentage quotas for each company were then converted into sales allocations on a tonnage basis for the world, the region and each national market.”\(^9\) This “budget” exercise was repeated in the fall of each year. In 1990, they recruited Eisai, Hoechst, Daiichi, E. Merck, and Takeda to cartelize a dozen other vitamins markets. By early 1991, most of the Roche-led cartels were successfully raising prices.

Regular meetings took place. At the annual “budget” meetings, the divisional heads of the vitamins business decided on the quotas and overall strategy. The managers of specific vitamins met four times a year to monitor the implementation of the quota systems. Regarding the monitoring structure, bulk-vitamins contracts are private and prices are not completely transparent, although firms learned general industry trends through customers.\(^10\) A more objective source was the customs clearance data in the international trade statistics. The ultimate confirmation of each member’s adherence to the quota relied on cross-validation of self-reported sales with trade statistics at the quarterly meetings. The geography of vitamins production facilitated the identification problem, because each firm operated only one or two plants for each vitamin, usually in its home country, and their nationality varied.\(^11\)

Throughout the 1990s, nothing like a full-blown price war was observed, except when several of the cartels ceased operations permanently. The firms did not formally specify contingency plans for adverse events, which is typical of cartel agreements in practice. Nevertheless, EC (2003) reveals they were occasionally expressing their shared belief that the

---

\(^7\) See Appendix A.1, “(a) The Nature and Reliability of the Evidence.”

\(^8\) See Appendix A.1, “(b) Competition before and after the Cartels.”

\(^9\) See Appendix A.1, “(c) Volume-Control Mechanism: ‘Budgets.’ ”

\(^10\) See Appendix A.1, “(d) Meetings” and “(e) Price Increases.”

\(^11\) See Appendix A.1, “(f) Monitoring” and “(g) World-Wide Markets and Production Locations.”
end of agreements would bring back competitive behaviors and low prices. Given the importance of off-equilibrium beliefs in the theory of collusion, we quote three such occasions in the vitamin C cartel: 12

- “Takeda complained that (...) if it did not get evidence that the European producers were following its price in May and June, it would ‘react’ against them” (recital 425);
- “According to BASF however, 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. Takeda relented and new lower vitamin C volume allocations were agreed among the four companies” (recital 442); and
- “Takeda returned to its favourite theme that it was ‘unreasonable to ensure the continuation of 1990 shares (...).’ Roche replied that if allotment cuts were mentioned to BASF and Merck they would stop following the scheme and bring chaos to the market with their low prices” (recital 444).

The 16 cartels ended in one of two ways. Six of them fell apart in 1994 or 1995, allegedly because of the entry and expansion by fringe producers and/or slowdown of demand growth. By contrast, the 10 other cartels were operating smoothly until prosecutions became imminent in the United States. In January 1999, RP applied for (and was granted) amnesty under the Department of Justice (DOJ)’s relatively untested Corporate Leniency Program. Within two months, both Roche and BASF pled guilty and agreed to pay record-breaking fines of $725 million in total. 13

The industrial landscape completely changed after 1999, as drug makers shed non-drug operations. RP’s application for amnesty coincided with its drugs division’s proposal to merge with Hoechst, a German rival. Regulators in America and Europe approved this merger, which created Aventis (currently Sanofi). Takeda sold its vitamin business to BASF in 2001. Roche sold its Vitamins and Fine Chemicals division to DSM, a Dutch chemicals firm, in 2003.

12 Curiously, we found no indication of contingency plans (“punishment”) connecting multiple vitamin markets, despite theoretical possibilities that such arrangements could have improved the ICC of each cartel (e.g., Bernheim and Whinston 1990). See Appendix A.1, “(h) (Lack of) Interaction between Different Cartels” and “(i) Adjustments.”
13 See Appendix A.1, “(j) Six ‘Natural Deaths’” and “(k) Ten ‘Antitrust Deaths.’”
2.2 Competitive Fringe in China

Chinese producers played a major role as the competitive fringe in the breakup of the vitamin C cartel and a few others. Their market entry and expansion were closely related to China’s economic liberalization in the 1990s, such as Deng Xiaoping’s official endorsement of private enterprises in 1992 and the liberalization of prices and commerce in 1994. Appendix A.2 summarizes key historical events and the profile of major producers in China.

In the mid-1980s, researchers at a government laboratory discovered and patented a new method to manufacture vitamin C at a lower cost and smaller scale than the existing technology. Nevertheless, “Chinese vitamin C production had been of little importance internationally” because “during the 1980s, exports were closely controlled by the state.”

Figure 2: Sudden Growth of Fringe Supply (Vitamin C)

Note: The dashed lines represent the cartel firms’ static predictions, as suggested by EC (2003).
Source: Bernheim (2002a) and EC (2003).

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.” Figure 2 shows the sudden growth of fringe supply in 1992, when the Bosnian war disrupted the Eastern European supply of vitamin C for the former socialist bloc (Li 2002). The government encouraged SOEs to increase production and supported the entry of private firms. More than 20 firms had entered by 1995.

---

None of these events seemed to be precisely forecasted by the cartel firms. EC (2003) reveals that even after three years of China’s rapid growth, the vitamin C 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 such “static” expectations). Underestimation and staggered revisions were common in other vitamin markets as well. Obtaining correct, timely information on the growth trajectory of fringe seems difficult. In section 6.2, we find the cartel could not have been incentive-compatible from the beginning if they had correctly foreseen the rapid growth of fringe.

2.3 Product Characteristics

We use the following basic features of the industry to guide our modeling choice. Each vitamin is used for its specific biochemical functions. Production of each vitamin required its own specific plant. Both UKCC (2001) and EC (2003) determined that “there was neither demand-side nor supply-side substitutability for individual vitamins.” The underlying production technologies of vitamins had matured by the 1980s and were common across firms. The only exception was the invention of a new method to produce vitamin C in China. No major entry or exit occurred during the cartel period except for fringe firms.

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. 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. Bernheim was engaged on behalf of more than 9,100 purchasers. EC (2003) recognizes a few “large” buyers, primarily in the vitamin C market, but even Coca-Cola, one of the world’s largest buyers, accounted for only 2.14% of the aggregate sales. Thus, buyer concentration is generally low.

Bulk vitamins are sold in different forms according to the product and the application: crystalline, in oil, with a protective coating, or in a powder matrix. Nevertheless, each cartel

---

16 See Appendix A.1, “(l) Repeated Failures to Predict the Growth of Fringe.” The cartels for vitamins B1, B6, and B9 repeatedly failed to achieve the “budgets” because of Chinese exports. The vitamin B2 cartel failed to predict the output growth of Coors/ADM in America.

17 See Appendix A.1, “(m) Health Benefits of Vitamins.” Other benefits for humans have not been proven.

18 See Appendix A.1, “(n) Substitutability between Vitamins,” and Appendix A.3 (production technology).

19 “Pink carotinoid,” or astaxanthin, was an exception. BASF was building a new plant and coordinating with Roche (the only producer of astaxanthin) for “controlled entry.” But the plant “did not come on stream until 1999” and the agreement was never implemented (EC 2003, recital 527).
used a single target price and tonnage-based sales allocation, by converting diluted products into the equivalent of 100% product. UKCC (2001) observes that customers shop around and switching is easy, because “as vitamins are chemical products, once they meet specified standards of purity, physical form, packaging and so on, it becomes difficult for a producer to differentiate its product from competitors.” Vitamin sales can be made either through long-term contracts or spot sales, but typical supply agreements last for only three months. Transportation costs and tariff barriers play minor roles: “The worldwide character of the markets” is confirmed by the internal organization of the cartels, as well as the presence of third-party arbitrage traders.

3 Data

3.1 Sources and Variables

As part of their investigation in 1999, the Federal Bureau of Investigation (FBI) and the DOJ obtained data on production and sales, personal records of cartel meetings, and individual depositions. These antitrust enforcements were followed by the purchasers’ multi-district litigation, which were consolidated at the U.S. District Court for the District of Columbia. Some of the expert reports (Bernheim 2002a, 2002b; Landes, Sider, and Bamberger 2002) were included in jury trials in 2003 and became publicly available.

Bernheim (2002a) relied primarily on the internal data from Roche, the cartel leader. He also had customer-specific transaction-level data from various plaintiffs and defendants. From his report, we build our dataset on the following variables between 1980 and 1998.

**Prices** Bulk vitamins are homogeneous goods within each category, but multiple grades of concentration exist. Roche’s dataset (ROVIS) contained “monthly weighted average prices (aggregated over customers) for a collection of vitamin products.” Following the cartel’s own practices, Bernheim converted them to “100% basis” and aggregated the product-level sales to a single price index.

**Production** Bernheim (2002a) shows aggregate outputs and each firm’s percentage market shares at the annual frequency, from which we calculate firm-level outputs. Similarly to

---

20 See Appendix A.1, “(o) Bulk Vitamins are Commodities,” and Appendix B.1 (firm-level prices).
21 See Appendix A.1, “(g) World-Wide Markets and Production Locations.”
22 A new round of antitrust cases emerged in the late 2000s and early 2010s when the Chinese firms, which had become dominant global players by then, restricted their export volumes (e.g., Bernheim 2008). We do not study this episode, because most of the main players in the 1990s had already exited, and the existence of export quotas (set by the Chinese Ministry of Commerce) complicated its legal characteristic.
the price indexes, all outputs are converted to 100% basis and aggregated at the vitamin level (i.e., the level at which the cartels defined and allocated quotas).

**Costs** Bernheim (2002a) computes unit costs by using the 100% basis price indices and the annual worldwide “contribution margins” (i.e., profit margins) for each vitamin family in the Roche Data Books. These costs are accounting measures based on engineering estimates, not economic marginal costs. Despite this limitation, we have chosen to use these unit-cost data for the following reasons. First, Roche’s unit-cost measure consists of labor, raw materials, and other intermediate inputs such as electricity (i.e., variable costs). Second, Roche used this measure for its own production decisions. Third, Roche and other firms’ unit costs seem constant at any quantity or “nameplate capacity utilization rates” (see Appendix B.2), which would limit the scope of divergence between average variable costs and marginal costs. Physical capacities at the facility level were far from binding during the sample period (see below), which would make the shadow price of capacity less of a concern.

**Table 1: Summary Statistics (Vitamin C)**

<table>
<thead>
<tr>
<th>Variable</th>
<th>Mean (and standard deviation)</th>
<th>Non-cartel period ($I_t = 0$)</th>
<th>Cartel period ($I_t = 1$)</th>
<th>Full sample</th>
</tr>
</thead>
<tbody>
<tr>
<td>Price (US$/kg)</td>
<td></td>
<td>11.21 (0.88)</td>
<td>13.61 (2.70)</td>
<td>12.43 (2.35)</td>
</tr>
<tr>
<td>Roche’s unit cost (US$/kg)</td>
<td></td>
<td>5.55 (0.60)</td>
<td>6.28 (0.64)</td>
<td>5.93 (0.71)</td>
</tr>
<tr>
<td>Aggregate output (1,000 MT)</td>
<td></td>
<td>44,570 (15,553)</td>
<td>51,393 (13,245)</td>
<td>48,161 (14,404)</td>
</tr>
<tr>
<td>of which Roche</td>
<td>20,967 (3,795)</td>
<td>21,744 (1,407)</td>
<td>21,376 (2,748)</td>
<td></td>
</tr>
<tr>
<td>of which Takeda</td>
<td>8,609 (2,942)</td>
<td>10,622 (3,092)</td>
<td>9,668 (3,113)</td>
<td></td>
</tr>
<tr>
<td>of which E. Merck</td>
<td>4,262 (438)</td>
<td>4,057 (621)</td>
<td>4,154 (537)</td>
<td></td>
</tr>
<tr>
<td>of which BASF</td>
<td>2,078 (1,497)</td>
<td>2,663 (745)</td>
<td>2,386 (1,167)</td>
<td></td>
</tr>
<tr>
<td>of which fringe</td>
<td>8,654 (8,141)</td>
<td>12,308 (9,443)</td>
<td>10,577 (8,807)</td>
<td></td>
</tr>
<tr>
<td>Number of months</td>
<td>112</td>
<td>116</td>
<td>228</td>
<td></td>
</tr>
</tbody>
</table>

*Note: The full sample period is January 1980–December 1998. The main cartel period is January 1991–August 1995. We also set $I_t = 1$ for an earlier (suspected) cartel episode between 1985 and 1988 as well as the 12 months after August 1995 to define the non-cartel period conservatively. Cost and output data are annual, and we treat 1985–1988 and 1991–1996 as the cartel period for these variables. MT (metric ton) is equal to 1,000 kg.*

**Capacities** Although Bernheim (2002a) had “nameplate capacity utilization” data on multiple firms at the vitamin level, he did not use them, because “nameplate capacity” corresponds to neither the long-run capacity (defined by the physical sizes of production facilities) nor the short-run capacity (in terms of weekly or monthly production schedules, such as work shifts and procurement plans for raw materials). Thus, we follow Bernheim and several other experts in not using this ambiguous measure, and interpret the actual

---

23Bernheim (2002a) 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).

24See Appendix B.3 for further details on capacities.
production as a reflection of the firms’ short-run production plans.


### 3.2 Data Patterns and Key Observations

Figure 3 (left) shows the worldwide production of vitamin C increased throughout the sample period. The four cartel members (Roche, Takeda, E. Merck, and BASF) restricted outputs between 1991 and 1995, but the competitive fringe in 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.

![Figure 3: Output, Price, and Cost (Vitamin C)](image)

*Note:* Shaded areas indicate cartel periods (January 1985–December 1988 and January 1991–August 1995). The first spell was not the target of the official investigation or litigation (and hence we do not have evidence on this suspected earlier episode); the second one corresponds to the main cartel episode.  
*Source:* Bernheim (2002a) and EC (2003).

Figure 4 shows similar pictures for vitamins A and E, and beta carotene. 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. A

---

25Bernheim (2002a) suggests earlier attempts of collusion in the mid-1980s and the price data take a year to reflect the cartel’s breakup (due to the lagged turnover of individual contracts).
A closer look at the output graphs reveals two differences across markets. One is the variation in market structure: Beta carotene is a duopoly, vitamin A is a triopoly, and vitamin E is a quadropoly. Another difference is the size of the competitive fringe. None existed in beta carotene; vitamins A and E had some small suppliers.

Figure 4: Output, Price, and Cost (Vitamins A and E, and Beta Carotene)
Two data patterns inform our subsequent analysis. First, both the price and the aggregate output increased between 1990 and 1994, which suggests a steady growth of demand. Second, the markup was sizeable even outside the cartel period (e.g., early 1980s), despite the experts’ view that bulk vitamins are commodities and the geographic market is global. Bertrand competition (with homogeneous goods, common technology, and multiple firms) would predict a much smaller or zero markup, whereas the Cournot model could reconcile these pieces of evidence. Hence, we use a quantity game to characterize the stage game.\textsuperscript{26}

4 Empirical Analysis of Demand and Costs

Sections 4 and 5 integrate the institutional details and the data in a structural model, and quantify the ICCs for collusion. Our exposition in the main text focuses on the vitamin C market, but we conduct the same analysis of the three other markets (i.e., vitamins A and E, and beta carotene) as well.

4.1 Demand

We specify a linear demand model for a bulk vitamin product,

\[ Q_t = \alpha_0 + \alpha_1 P_t + \alpha_2 X_t + \varepsilon_t, \]

where \( P_t \) is the “100% basis” price index, \( X_t \) is a collection of demand shifters, and \( \varepsilon_t \) is a non-systematic component of demand shifters.\textsuperscript{27} A usual procedure for demand estimation would be to (i) collect data on \( X_t \), (ii) construct instrumental variables (IVs) for \( P_t \), and (iii) perform an IV regression of \( Q_t \) on \( P_t \) and \( X_t \) to estimate \( (\alpha_0, \alpha_1, \alpha_2) \). Supply-side “shocks” in our data, such as Roche’s unit cost \( \left(c_{roche,t}\right) \) and the indicator of cartel periods \( (I_t) \), are natural candidates for IVs.

However, this approach faces three practical problems in the current context. First, not all of \( X_t \) is observable. Data on population and income of humans (e.g., GDP) and the number of slaughtered animals exist, but the perceived health benefits of vitamins are intangible. Second, the demand for most vitamins was growing steadily. The simultaneous increase of \( Q_t \) and \( P_t \) leads to a scatter plot of \( (Q_t, P_t) \) with an upward-sloping pattern, and creates severe multicollinearity (see Appendix C.1). Third, the supply-side IVs lack sufficient

\textsuperscript{26}We revisit this modeling choice in section 6.1.

\textsuperscript{27}We also consider log-linear specifications in section 6.2.
variation to overcome this multicollinearity problem, because $c_{i,t}$ is mostly constant over time, and the regime change in $I_t$ is infrequent. For these reasons, we found IV regressions lead to “upward-sloping” demand-curve estimates (i.e., $\hat{\alpha}_1 > 0$), unless we make highly ad hoc functional-form assumptions or select a special subsample (e.g., 1995–1997, during which the price decreased sharply as the cartel collapsed).

We adopt an alternative identification strategy that exploits supply-side restrictions. First, we assume a static Nash equilibrium of the quantity game in the competitive periods ($I_t = 0$). Roche’s first-order condition (FOC) is

$$P_t + \frac{dP}{dQ} \times q_{roche,t} = c_{roche,t} \quad \text{if } I_t = 0,$$

(2)

where $q_{roche,t}$ is its output. We observe $(P_t, q_{roche,t}, c_{roche,t})$ in the data, so $\frac{dP}{dQ}$ is the only unknown. Note the slope of the demand curve is simply the reciprocal of $\frac{dP}{dQ}$: $\alpha_1 = (dP/dQ)^{-1}$. Hence, $\alpha_1$ is identified whenever $\frac{dP}{dQ}$ is. Rearranging the FOC, we obtain the estimating equation,

$$P_t - c_{roche,t} = \frac{dP}{dQ} \times (-q_{roche,t}) + \eta_{roche,t} \quad \text{if } I_t = 0,$$

(3)

where we can regress $(P_t - c_{roche,t})$ on $(-q_{roche,t})$ to estimate $\frac{dP}{dQ}$, and $\eta_{roche,t}$ is the error term to account for potential measurement errors in $c_{roche,t}$.

Second, we do not use data from the collusive periods ($I_t = 1$) for estimation purposes. We apply the same constant $\hat{\alpha}_1$ based on its estimates from the non-cartel periods ($I_t = 0$) throughout the sample period.

Third, for the entire sample period, we obtain the intercept of the demand curve in each $t$ by subtracting $\hat{\alpha}_1 P_t$ from $Q_t$ in (1) to calculate the effective demand shifter/intercept,

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

(4)

as the residual. We cannot identify each term separately, but we need not distinguish between them for our subsequent analysis.

---

28 Similar methods are used in merger investigations by antitrust agencies, according to Nathan Miller of Georgetown University. See Genesove and Mullin (1998) and MacKay and Miller (2018) for related studies.
4.2 Costs

We use Roche’s cost data to identify demand. Bernheim (2002a) contains data on a few other firms’ costs as well, but they are incomplete.\(^ {29} \) Accordingly, we have chosen to estimate their marginal costs by using their respective FOCs,

\[
P_t + \frac{dP}{dQ} \times q_{i,t} = c_{i,t} \quad (i \neq \text{roche}) \quad \text{if } I_t = 0,
\]

where \( c_{i,t} \) is the unit cost at non-Roche firm \( i \). Given their relatively stable market shares, we assume constant gaps exist between their costs and Roche’s,

\[
c_{i,t} = c_{\text{roche},t} + \gamma_i \quad (i \neq \text{roche}),
\]

where \( \gamma_i \) is the mean difference between \( c_{i,t} \) and \( c_{\text{roche},t} \).\(^ {30} \)

Rearranging (5) and (6), the estimating equations for non-Roche firms are

\[
P_t - c_{\text{roche},t} = \gamma_i + \frac{dP}{dQ} \times (-q_{i,t}) + \eta_{i,t} \quad (i \neq \text{roche}) \quad \text{if } I_t = 0,
\]

where \( \eta_{i,t} \) is the error term to account for firm-specific measurement errors in \( c_{i,t} \). We can jointly estimate \( \frac{dP}{dQ} \) and \( \gamma_i \) by minimizing the sum of squared residuals in (3) and (7),

\[
\sum_t \sum_i \eta_{i,t}^2 \quad (i = \text{roche, takeda, e.merck, basf}) \quad \text{if } I_t = 0.
\]

4.3 Estimates of Demand and Costs

An additional complication to our (otherwise straightforward) econometric problem is that \( P_t \) is recorded at a monthly frequency, whereas \( c_{i,t} \) and \( q_{i,t} \) are observed only annually. The coarseness of \( c_{i,t} \) poses a relatively minor problem, because Roche’s cost was stable over 19 years (see Figure 3), and it is on the LHS of the regression equations, which means the error term accommodates its measurement error. Hence, we use the same (annual) observation of \( c_{\text{roche},t} \) within each year. By contrast, measurement error in \( q_{i,t} \) requires a more thorough treatment, because outputs might vary substantially across months, and they belong to the RHS of the estimating equations, which potentially creates an omitted-variable bias.

Table 2 reports two versions of parameter estimates. Column 1 uses the simplest (if \( \text{ad hoc} \) approach of imputing monthly \( q_{i,t} \) by dividing each annual observation \( q_{i,y(t)} \) by

\(^{29}\)See Appendix C.3 for non-Roche firms’ cost data and our estimates.

\(^{30}\)An alternative specification with constant ratios (i.e., \( c_{i,t} = c_{\text{roche},t} \times \gamma_i \)) generates similar results.
Table 2: Estimates of Demand and Costs (Vitamin C)

<table>
<thead>
<tr>
<th>Imputation method</th>
<th>Equal split</th>
<th>EM algorithm</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \frac{dP}{dQ} )</td>
<td>-2.829 \quad (0.107)</td>
<td>-2.817 \quad (0.110)</td>
</tr>
<tr>
<td>( \gamma_{\text{takeda}} )</td>
<td>3.356 \quad (0.060)</td>
<td>3.360 \quad (0.058)</td>
</tr>
<tr>
<td>( \gamma_{\text{e-merck}} )</td>
<td>4.550 \quad (0.045)</td>
<td>4.554 \quad (0.045)</td>
</tr>
<tr>
<td>( \gamma_{\text{basf}} )</td>
<td>4.930 \quad (0.084)</td>
<td>4.935 \quad (0.084)</td>
</tr>
<tr>
<td>Number of observations (months)</td>
<td>112</td>
<td>112</td>
</tr>
</tbody>
</table>

Note: Standard errors in parentheses are based on 1,000 block-bootstrap samples, where each block consists of 12 consecutive months of a calendar year. In column 1, we divide each annual observation of firm-level output by 12 to impute monthly outputs. In column 2, we use the EM algorithm to impute monthly outputs as part of the estimation procedure (see Appendix C.2 for details). See Appendix C.5 for vitamins A and E, and beta carotene.

12, where \( y(t) \) denotes the year to which month \( t \) belongs. We then take these imputed \( \tilde{q}_{i,t} \) as data and minimize (8) to estimate \( \frac{dP}{dQ} \) and \( \gamma_i \) (for all \( i \neq \text{roche} \)). In column 2, we discipline the imputation of \( \tilde{q}_{i,t} \) by using the expectation-maximization (EM) algorithm as part of the estimation procedure. We set the initial values of \( \tilde{q}_{i,t} \) as in column 1. The first or “maximization” step takes these initial “data” as given and minimizes (8) to estimate the parameters. In turn, the second or “expectation” step takes these parameter values as given, and estimates the conditional expectation of \( \tilde{q}_{i,t} \) under the accounting constraint that monthly outputs must sum to the annual data for each firm in each year. We repeat these two steps until convergence (see Appendix C.2 for details). Regardless of the imputation methods, the coefficient estimates are similar and within the standard errors of each other. We use 1,000 block-bootstrap samples to calculate standard errors, where each block consists of 12 consecutive months of a calendar year.

Figure 5 shows the implied price elasticity of demand, \( \hat{e}_t \equiv \frac{\partial Q_t}{\partial P_t} \times \frac{P_t}{Q_t} \), and \( \tilde{X}_t \) along the historical trajectory of \( (Q_t, P_t) \). The elasticity was relatively stable in the elastic region of the demand during the cartel period. The demand shifter/intercept grows steadily but slows down after 1995, which agrees with the documented evidence.\(^{31}\)

\(^{31}\)EC (2003, recital 454) and Bernheim (2002a, Appendix E).
Figure 5: Elasticity and Growth of Demand (Vitamin C)

Note: The dashed lines reflect standard errors based on 1,000 block-bootstrap samples.

4.4 Stage-Game Profits

We calculate each firm’s stage-game profits for all $t$,

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

in three different cases: $(\pi_{i,t}^C, \pi_{i,t}^D, \pi_{i,t}^N)$. Note we assume a constant marginal cost with respect to output level, because our data on cost, output, and “nameplate capacity” exhibit such patterns (see Appendix B.2). Let $\pi_{i,t}^C$ denote the profit under the cartel agreement. We assume the cartel’s collective output $Q_{car,t}$ maximizes its total profit,

$$\sum_{i \in I} \left( \frac{dP_t}{dQ_t} \times \left( Q_{car,t} + Q_{frist,t} - \tilde{X}_t \right) - c_{i*,t} \right) \times q_{i,t},$$

where $I = \{1, ..., n\}$ denotes the fixed set of cartel firms and $i^*$ is the cartel leader (Roche, which is also the lowest-cost producer in most vitamin markets). This assumption is based on our finding that the cartel did achieve high prices that are statistically indistinguishable from theoretical monopoly prices (see Figure 6 below).

Regarding the individual output levels, the cartel agreement used percentage quotas based on the member firms’ respective within-cartel market shares in the pre-cartel period (e.g., 1990 in the vitamin C cartel).\footnote{Appendix B.4 shows within-cartel market shares are stable over time.} These percentage quotas were applied to the collective
Figure 6: Actual and Monopoly Prices (Vitamin C)

Note: The hypothetical monopoly price is based on the demand estimates (net of fringe supply) and Roche's cost data. The thin dashed lines indicate standard errors based on 1,000 block-bootstrap samples.

output $Q_{car,t}$ and converted into sales allocations on a 100%-equivalent tonnage basis, $(\bar{q}_{i,t})_{i \in I}$. Accordingly, we assume the cartel firms are supposed to supply their respective quotas and earn $\pi^C_{i,t} = (P_t - c_{i,t}) \bar{q}_{i,t}$.

Likewise, $\pi^D_{i,t}$ is an individual member’s optimal non-compliance profit against $(\bar{q}_{j,t})_{j \neq i}$. That is, a cartel firm can maximize its individual profit by secretly increasing its own output beyond its allotment ($q_{i,t} > \bar{q}_{i,t}$) while the other members stick with their quotas. Finally, $\pi^N_{i,t}$ is the static Nash equilibrium profit.

Figure 6 compares the historical price data and the theoretical cartel/monopoly price that corresponds to $\pi^C_{i,t}$ (given the actual demand, fringe supply, and cost profile at each $t$). The actual price is low at the beginning of the cartel in 1991, which is not surprising, because our demand estimation assumed the spot market is in Nash equilibrium until December 1990. Subsequently, the actual price converges to the optimal cartel price by the spring of 1993, which is surprising because we do not impose monopoly pricing (or any other restrictions) when $I_t = 1$. This pattern agrees with the evidence in EC (2003, recital 432) that the vitamin C cartel achieved its price target in that year and switched to less aggressive price increases.\(^{33}\)

The cartel took more than a year to achieve its optimal price. EC (2003) shows the

---

\(^{33}\)The actual and monopoly prices of vitamins A and E, and beta carotene show similar patterns.
firms took turns announcing small price increases, in an attempt to mimic “natural” price-leadership behaviors. Note the data reflect the average of all transactions at Roche. Although typical contracts spanned three months, some lasted for six to 12 months. Hence, after the last documented cartel meeting on August 24, 1995, the price data also took a year to come down to the pre-cartel level. We do not model these sluggish transitions that are driven by such idiosyncratic and mechanical features, and instead focus on more fundamental components of the ICCs: $(\pi_{i,t}^C, \pi_{i,t}^D, \pi_{i,t}^N)$.

5 Dynamic Implications on Cartel Stability

This section studies the dynamic implications of the (static) estimates of demand and costs in the previous section. Section 5.1 presents our repeated-games model, which we specify according to the internal organization of the vitamin cartels in EC (2003). Section 5.2 uses this dynamic model to estimate the ICCs. Section 5.3 assesses the effects of demand slowdown and fringe supply to explain the collapse of the vitamin C cartel in 1995.

5.1 A Repeated-Games Framework

We introduce the following repeated-games framework to quantify the ICC for collusion, which are necessary for assessing the impacts of various market forces on the stability of collusion, including the coordinated effects of a merger. In each month $t$, each cartel firm $i \in I$ observes the demand (both its slope $\frac{dP}{dQ}$ and effective intercept $\bar{X}_t$) and the cost profile $(c_{j,t})_{j \in I}$, and decides its output $q_{i,t}$ conditional on its private history $h^i_t$ (to be defined). Their total supply $Q_{car,t} = \sum_{i \in I} q_{i,t}$, fringe supply $Q_{fri,t}$, and the demand curve collectively determine the market price $P_t = \frac{dP}{dQ} \times \left( Q_{car,t} + Q_{fri,t} - \bar{X}_t \right)$.

Based on the monitoring structure of the vitamin cartels in section 2.1, we assume the quantity profile is observed with $L$ periods of lag: $\left((q_{j,\tau})_{j \in I}, Q_{fri,\tau}\right)_{\tau \leq t-L}$. We set $L = 3$ because most cartels relied on government statistics to verify self-reports at quarterly meetings. Under this institutional setup, firm $i$’s private history at the beginning of month $t$ is

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

Our results are robust to the consideration of long-term contracts (see section 6.2). Harrington and Chen (2006) propose a model of an endogenously rising price path.

We abstract from the exact timing of these quarterly meetings on calendar due to data limitations.
and the public history is

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

Regarding the cartel firms’ beliefs, we assume they held static expectations about fringe output (i.e., they expect $Q_{fri,\tau} = Q_{fri,t-3}$ for each $\tau \geq t$), because section 2.2 suggests they had limited visibility about the growth of fringe supply. By contrast, the industry background in section 2.3 suggests the evolution of demand and costs, $\left(\hat{X}_{\tau}, \frac{dP}{dQ}, (c_{j,\tau})_{j \in I}\right)_{\tau=1}^{\infty}$, is common knowledge, and hence we assume rational expectations and omit them from the history. Finally, we assume the demand, the costs, and fringe supply after the sample period remain constant at their end-of-sample values in December 1998.

We consider the following equilibrium based on trigger strategies. The cartel firms are supposed to supply their respective quota allocations, $(\tilde{q}_{i,\tau}|t)$, given the rational expectation of demand and costs $\left(\hat{X}_{\tau}, \frac{dP}{dQ}, (c_{j,\tau})_{j \in I}\right)_{\tau \geq t}$ and the static expectation of fringe output $Q_{fri,\tau} = Q_{fri,t-3}$ for each $\tau \geq t$. 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 current period $t$.

We say non-compliance is confirmed in period $\tau$ if, given the government statistics, it becomes common knowledge that some firm did not produce $q_{i,s}|t$ for the first time (i.e., $q_{i,s}|t = q_{i,s}|t$ for each $s < \tau - 3$ and $i \in I$, but $q_{i,\tau-3}|t \neq q_{i,\tau-3}|t$ for some $i \in I$).

In each month $t$, given the expectation formed at $t$, the firms agree to play the following strategy for $\tau \geq t$: (i) if no non-compliance is confirmed before month $\tau$, then each firm sells $q_{i,\tau}|t = q_{i,\tau}|t$; and (ii) if some non-compliance is confirmed in some previous month $s \leq \tau$, then each firm sells a static Nash equilibrium quantity $q_{i,\tau} = q_{i,\tau}|t$. Let us call this strategy the “trigger strategy.”\(^36\)

Complying with the cartel agreement from month $\tau$ on gives firm $i$ the payoff of

$$V_{i,\tau}|t = \sum_{s \geq \tau} \beta^{s-\tau} \pi_{i,s}|t,$$

where $\beta \in (0, 1)$ is the discount factor. When firm $i$ does not comply at $\tau$, the optimal

\(^36\)Note that producing $q_{i,t}^N$ in each period $t$ is an equilibrium in the dynamic game as well, because the evolution of demand and costs does not depend on the firms’ outputs.
deviation payoff is

\[ V_{i,t}^D = \sum_{\tau=1}^{\tau+2} \beta^{s-\tau} \pi_{i,s,|t}^D + \sum_{s=\tau+3} \beta^{s-\tau} \pi_{i,s,|t}^N, \]

(12)
because no punishment is conducted until the government statistics verify \( i \)'s non-compliance \( L = 3 \) months later.

In each \( t \), given \( V_{i,t}^C \) and \( V_{i,t}^D \) for \( \tau \geq t \), if there exist \( i \in \mathcal{I} \) and \( \tau \geq t \) for which \( V_{i,t}^C < V_{i,t}^D \), it becomes common knowledge as of \( t \) that some firm will deviate in month \( \tau \) and the prevailing actions will be a static Nash equilibrium from \( \tau + 3 \). The situation becomes the same as in a finitely repeated game, and hence each firm deviates in month \( t \). Accordingly, the trigger strategy is an equilibrium (given the firms’ expectations at \( t \)) if and only if

\[ \min_{i \in \mathcal{I}, \tau \geq t} (V_{i,t}^C - V_{i,t}^D) \geq 0. \]

(13)

That is, the cartel’s stability at \( t \) requires its members to expect positive payoffs at each point in the foreseeable future \( \tau > t \) given the information at \( t \).

Recall that the cartel firms held static beliefs about \( Q_{fr,\tau} \). The actual \( Q_{fr,t} \) in the data increased from 1992 in a staggered manner in the vitamin C market, which is the only case (among the four cases we scrutinize) in which the cartel collapsed before the prosecution. The firms revise their expectation of \( Q_{fr,\tau} \) more pessimistically over time. Eventually, the cartel becomes unstable in the period in which (13) is violated for the first time.

This setup implies that, 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, the cartel firms 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 \( Q_{fr,t} \) 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.
5.2 Estimates of the ICCs

Figure 7 shows our ICC estimates for the vitamin C cartel (i.e., the LHS of equation 13) as well as its individual members. Roche’s ICC was the least stringent of all members’ at the beginning of the cartel, because it is the largest firm with the lowest cost, and the cartel quotas are well aligned with market shares under static Nash (recall quotas are based on competitive market shares in 1990). However, Roche’s ICC became more stringent as the negative news on the fringe supply arrived, and was the tightest in 1995 and 1996, coinciding with the cartel’s breakup in reality.

The finding that the cartel was destabilized by the sudden, unexpected growth of fringe would seem obvious, but note the ICC would never be violated if costs were symmetric. Larger $Q_{fri,t}$ means smaller residual demand, $\bar{X}_t - Q_{fri,t}$. When residual demand is small, a low-cost firm such as Roche could win a disproportionately large market share (relative to high-cost rivals) in static Nash,\(^{37}\) but the quota share is fixed. Hence, low-cost firms become increasingly disgruntled, as the residual demand shrinks. We found any levels of annual discounting $\beta \in [0.75, 0.95]$ leads our model to predict high probabilities of the cartel’s collapse in 1995 and 1996,\(^{38}\) primarily as a result of such growing discrepancies between the status quo and the initial situation.

By contrast, the same analysis yields completely different pictures for vitamins A and E, and beta carotene. Figure 8 shows Roche’s ICCs are comfortably above zero, suggesting the stability of these cartels throughout the 1990s. In reality, the three cartels were fully operational until the prosecution in 1999. Hence, our model explains the life and death of all four cartels for which reliable data exist.

5.3 Why the Vitamin C Cartel Collapsed

Why did the vitamin C cartel collapse within a few years, whereas the other three survived for a decade? Figure 9 illustrates the effects of demand growth and fringe supply on the

\(^{37}\)To verify this claim, suppose the inverse demand function is $P(Q) = a - bQ$. Firm $i$’s market share in Cournot equilibrium is

$$s_i = \frac{a - (n + 1)c_i + \sum_j c_j}{na - \sum_j c_j},$$

where $c_i$ is firm $i$’s marginal cost. As $a \to \infty$, $s_i$ converges to $\frac{1}{n}$. Hence, cost asymmetry matters less when demand is larger.

\(^{38}\)Appendix C.4 shows Roche’s ICC under different $\beta$. Note our notion of $\beta$ includes the probability of an “exogenous” death of the cartel, and the main decision-makers are divisional heads, not CEOs or shareholders.
Figure 7: Collective and Individual ICCs (Vitamin C)

Note: We set $\beta = 0.85$ for these plots. The “90% confidence intervals” reflect the 95th and 5th percentiles of the 1,000 block-bootstrap estimates. All values are rescaled as the average monthly profits (i.e., multiplied by $1 - \beta$). See Appendix C.4 for estimates under different levels of $\beta$. 

24
Figure 8: Roche’s ICCs (Vitamins A and E, and Beta Carotene)

Note: We set $\beta = 0.85$ for these plots. The “90% confidence intervals” reflect the 95th and 5th percentiles of the 1,000 block-bootstrap estimates. All values are rescaled as the average monthly profits (i.e., multiplied by $1 - \beta$).

cartel’s stability, by showing Roche’s ICC in four different scenarios.

Scenario 1 at the top is the “dream world” counterfactual for the cartel, in which the demand shifter did not decline after 1994, and the Chinese exports stopped growing after 1994 (i.e., $\tilde{X}_t^{CF1} = \tilde{X}_{1994}$ and $Q_{fri,t}^{CF1} = Q_{fri,1994}$ for $t > 1994$). Under these favorable conditions, Roche’s ICC in 1995 would have been greater than that in reality as of 1991. The cartel could have survived throughout the sample period.

Scenario 2 is the same as Scenario 1 on the demand side but incorporates the actual path of fringe supply (i.e., $\tilde{X}_t^{CF2} = \tilde{X}_{1994}$ and $Q_{fri,t}^{CF2} = Q_{fri,1994}$ for $t > 1994$). Hence, their difference reflects the impact of fringe growth in the mid-1990s. Even the lowest part of the ICC in this scenario is greater than the actual ICC in 1991. Thus, despite the rapid expansion of the competitive fringe, the cartel could have survived beyond 1995 if the market size had not shrunk at the same time.

Scenario 3 is the opposite of Scenario 2, incorporating the slowdown of demand while “freezing” the growth of fringe supply (i.e., $\tilde{X}_t^{CF3} = \tilde{X}_{1994}$ and $Q_{fri,t}^{CF3} = Q_{fri,1994}$ for $t > 1994$). Hence, we may interpret the difference between Scenarios 1 and 3 as the impact of the demand slowdown after 1994. Roche’s ICC in 1995 is still positive but as low as its actual level in 1991.39

Finally, the bottom ICC is the same one as in Figure 7. It reflects both the slowdown of the demand growth and the final phase of the Chinese supply expansion in the mid-1990s. Thus, our counterfactual simulations suggest neither the slowdown of demand growth nor the sudden expansion of fringe supply alone could explain the cartel’s breakup, but the two

Figure 9: Effects of Demand and Supply on Roche’s ICC (Vitamin C)

Note: We set $\beta = 0.85$ for expositional purposes, but similar patterns emerge under different values of $\beta$.

together could.

These simulations demonstrate how one can use our approach to measure the effects of various factors on cartel stability. In section 7, we extend this approach to measuring the effects of mergers (i.e., changes in market structure and cost profile) as well. Before proceeding, however, we discuss several theoretical considerations and conduct sensitivity analysis on the key features of the model.

6 Theoretical Considerations and Sensitivity Analysis

Section 5 showed a simple repeated-games model could explain the life and death of the four vitamin cartels. Section 6.1 considers more complicated theoretical models and discuss the implications of our modeling choices. Section 6.2 reports the results under alternative assumptions.

6.1 Theoretical Considerations

We discuss five important aspects of the model and their background.

Multi-market Contact Multi-market contact (MMC) is a situation in which firms compete (or collude) in more than one market. Theoretically, MMC could facilitate collusion if
firms adopt a scheme that involves punishment in multiple markets, thereby making punishment more severe and the ICC less stringent than in the single-market scheme. Curiously, the vitamin cartels did not seem to employ such punishment schemes, nor did the breakup of a cartel trigger punishment in other markets, even when the same members (and the cheating firms) were involved. Accordingly, we model punishment in each market separately, and our ICC estimates represent a lower bound of theoretically possible ICCs under multi-market punishment schemes.

Quantity Game We model the stage game as a quantity game, which is a standard model for homogeneous-good industries such as bulk chemicals. A usual criticism of the Cournot model is that real-world oligopolists choose both quantities and prices. We have considered alternative specifications including a price game (Bertrand) and a quantity-and-price game (Kreps-Scheinkman). Kreps and Scheinkman (1983) offer a “micro foundation” of Cournot competition, although some details require further discussion in the context of repeated games. Appendix D.1 considers these alternatives.

Delayed Monitoring In our model, no punishment is conducted until non-compliance is verified with $L$-month lag. That is, we assume “perfect” monitoring of sales quantities with a delay, based on the fact that the vitamin cartels relied on government statistics as a source of external verification of self-reported sales, at the quarterly meetings. Two questions arise. First, why did firms choose to wait for the government statistics when they might have been able to infer rivals’ non-compliance from contemporaneous prices and other (albeit imperfect) monitoring devices? Second, why did firms bother to exchange self-reported sales records when more reliable information would arrive in a few months anyway? Appendix D.2 investigates these issues.

Static Expectations on Fringe Supply We assume the cartel firms had static expectations about fringe supply, based on EC (2003). Were they really so “irrational”? One possibility is that they truly lacked precise information about Chinese SOEs and other fringe suppliers, which seems plausible given their record of failed predictions in multiple markets. (As we find in section 6.2, the cartel could not have even started operations if they had correctly foreseen the growth of fringe.) Another possibility is that they obtained correct information but suffered “behavioral” biases such as overconfidence. Yet another possibility is that they had correct predictions but rationally chose to speak naïvely at the cartel meet-

---

40 Bernheim and Whinston (1990), Matsushima (2001), Kobayashi and Ohta (2012), and Sekiguchi (2015).
41 See Appendix A.1, “(Lack of) Interaction between Different Cartels.” Why did they not use multi-market punishment? Our results in section 4.4 suggest the initial ICCs were satisfied even under the single-market punishment scheme, which obviates the need for more complicated arrangements. Coordinating 21 firms across 16 markets would pose a practical challenge as well.

**Nash Reversion** We specify an infinite repetition of the static Nash equilibrium as the punishment, because EC (2003) reports the cartel members communicated their understanding that someone’s withdrawal or non-compliance would bring back competitive behaviors and low prices (without any indication of reverting back to cooperative behaviors subsequently). In Appendix D.4, we consider other, more severe forms of punishment that have been proposed in the theoretical literature.

### 6.2 Sensitivity Analysis

We used as much direct evidence as possible to guide our modeling choices. This subsection considers alternative specifications of the following seven aspects.

**1) Endogenous Fringe** China was a major disturbance to cartel stability. Our baseline model treated its government and SOEs as exogenous fringe players who built new plants for geopolitical reasons and dumped domestic surpluses on the international market. In Appendix E.1, we model China as “endogenous” fringe that responds to the international price, by specifying and estimating a fringe-supply curve.\(^{42}\)

**2) Renegotiating Quotas** We relied on direct evidence from EC (2003) to model the cartel’s quotas based on its members’ pre-cartel market shares (i.e., under oligopolistic competition). The record indicates that whenever a member proposed a revision of pre-existing quotas, the “Big Three” opposed the idea because they believed negotiations over quotas would never end and would destabilize cartel operations. Such pragmatic considerations notwithstanding, we ask whether the vitamin C cartel could have optimally renegotiated quotas to avoid its collapse in 1995. In Appendix E.2, we find some quota reallocation could have relaxed Roche’s ICC sufficiently but would have been impractical.

**3) Cartel Price below Monopoly Level** Section 4 showed the actual cartel pricing converged to theoretical monopoly levels. Accordingly, our subsequent analysis equated cartel prices with monopoly prices. Appendix E.3 shows alternative cartel prices could not have relaxed the ICC.

**4) Monitoring Lag** Given the prominent role of monitoring in the theory of repeated games, one may naturally ask how the results change under alternative specifications of

\(^{42}\)We also consider the theoretical possibility that the cartel had known this 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 sustainable.
monitoring lag, \( L \). Appendix E.4 shows two cases, with \( L = 1 \) and \( L = 12 \). We find \( L = 1 \) would relax the ICC but only slightly, whereas \( L = 12 \) would make collusion unsustainable.

(5) **Rational, Static, and Adaptive Expectations** Our baseline model assumes the cartel members had rational expectations about demand and costs, as well as static expectations on China, based on the evidence in section 2. Appendix E.5 considers two alternatives. First, we use adaptive expectations on demand growth \((\tilde{X}_t)\), in which firms update predictions based on their observations in the recent past. Second, what if the firms had had rational expectations on China? In both cases, we find the ICC would have been violated from the beginning of the cartel (i.e., \( ICC_{1991} < 0 \)), which is inconsistent with the cartel’s existence for more than four years.

(6) **Log-Linear Demand** We used the linear demand model in section 4, whereas Appendix E.6 uses the log-linear specification. Neither the identification strategy nor the estimates of the price coefficient/elasticity and costs changes materially. However, the implied demand shifter \((\tilde{X}_t)\) exhibits patterns that disagree with the industry experts’ testimony, which makes us prefer the linear specification.

(7) **Long-term Contracts** A typical contract is either spot or lasts for three months, but some last six to 12 months (UKCC 2001). Our baseline model assumes monthly spot-market transactions and abstracts from multi-period aspects. Appendix E.7 shows explicitly incorporating long-term contracts does not alter the ICC.

7 **Mergers, Asymmetry, and Coordinated Effects**

Mergers affect the incentives for collusion, explicit or tacit.\(^{43}\) “Coordinated effects” refer to this possibility and have been an important issue in antitrust policy. Nearly 60% of challenges filed by the DOJ and FTC over 1990–2014 allege coordinated effects, according to Gilbert and Greene (2015). The central question is whether a merger makes firms more likely to restrict output and raise prices collectively, or (if the market currently has effective existing coordination) whether “it leads rivals to coordinate their strategic choices more perfectly, more completely, or more durably.”\(^{44}\)

Because the impact of a merger on collusive incentives involves multiple, possibly conflicting factors, the report to the EC’s Directorate General of Competition (DG-Comp) by

\(^{43}\) We regard our analysis of explicit collusion as a benchmark and an upper bound of what tacit colluders would be able to achieve. Theoretical models usually do not distinguish between explicit and tacit collusion, but see Awaya and Krishna (2016) for one that distinguishes between them. Kaplow (2013) elaborates on exactly what constitutes price fixing in the legal sense.

\(^{44}\) See Dick (2003, p. 66), then acting chief of competition policy at the DOJ’s Antitrust Division.
Ivaldi, Jullien, Rey, Seabright, and Tirole (2003) recommends a structural analysis “to develop a clear understanding of why each dimension is relevant, as well as of how it affects collusion — and is affected by a merger.” Our empirical model makes such coordinated-effects analysis possible. Because we use only such basic data as prices, costs, quantities, and the state of competition, which are often available to the antitrust authority at the time of merger review, our approach does not add extra burden relative to the data requirement for a conventional analysis (e.g., prospective merger simulations of unilateral effects).

In section 7.1, we explain how the ICC depends not only on the number of firms but also on their cost asymmetry. We then present empirical measures of coordinated effects in sections 7.2 through 7.4 by simulating ICCs under a series of hypothetical mergers. Results suggest many mergers could make market structure less competitive (in terms of unilateral effects) and less collusive (in terms of coordinated effects) at the same time.

7.1 How Mergers Affect the ICC

The conventional wisdom since Stigler (1964) is that mergers facilitate collusion, but this conjecture is not always correct. Rothchild (1999) points out that “the stability of a cartel depends not only on the number of members, but also on their individual or aggregate costs.” Subsequent theoretical works by Compte, Jenny, and Rey (2002), and Vasconcelos (2005) suggest mergers may hurt collusion if firms become more asymmetric after the deal. In particular, Vasconcelos (2005) highlights the role of efficiency gains in reshaping cost asymmetries among firms, which could either help or hinder collusion.45

Let \(s_i\) be firm \(i\)'s market share in the static Nash equilibrium (we omit the time subscript and fringe output for expositional purposes). Its profit under the quota-based cartel is

\[
\pi_i^C = s_i (P^m - c_i) Q^m,
\]

where \(P^m\) and \(Q^m\) are monopoly price and output based on the cartel leader's cost. Given the inverse demand function \(P = a - bQ\), the deviation payoff is

\[
\pi_i^D = \max_{q_i} (a - b (q_i + (1 - s_i) Q^m) - c_i) q_i.
\]

---

45 Vasconcelos (2005) considers optimal punishment, and most of his “counter-intuitive” findings come from easing the punishment. Hence, his results are not directly applicable to our context with Nash reversion. For this reason, we present our version of theoretical analysis in the following.
The punishment (static Nash) payoff is
\[\pi_i^N = s_i \left( P^N - c_i \right) Q^N,\]
where \( P^N \) and \( Q^N \) are the price and total output in the static Nash equilibrium. Firm \( i \)'s gains from collusion is proportional to
\[\pi_i^C - \pi_i^N = s_i \left[ (P^m - c_i) Q^m - (P^N - c_i) Q^N \right],\] (14)
and its opportunity cost (i.e., the deviation gain) is proportional to
\[\pi_i^D - \pi_i^C = \max_{q_i} \left( a - b (q_i + (1 - s_i) Q^m) - c_i \right) q_i - s_i \left( P^m - c_i \right) Q^m.\] (15)

Mergers reduce the number of firms and affect the ICC through four channels. First, fewer competitors mean a higher market share, \( s_i \). This direct market-share effect increases the gains from collusion (14) and relaxes the ICC.

Second, the gains from deviation (15) are decreasing in \( s_i \):\[
\begin{align*}
\frac{d}{ds_i} (a - b (q_i + (1 - s_i) Q^m) - c_i) q_i^* - s_i \left( P^m - c_i \right) Q^m &= b Q^m q_i^* - (P^m - c_i) Q^m \\
&= - (P^m - c_i - b q_i^*) Q^m < 0,
\end{align*}
\]
where \( q_i^* \) is firm \( i \)'s optimal unilateral deviation from the cartel agreement (i.e., best response to the other firms that are collectively producing \( (1 - s_i) Q^m \)). The first equality follows from the envelope theorem. The last inequality follows from firm \( i \)'s profit maximization, because its optimal unilateral deviation, \( q_i^* \), should never push the price below its own cost: \( P^m - b q_i^* > c_i \). Hence, \( \pi_i^C \) increases more than \( \pi_i^D \) does, as \( s_i \) increases with mergers. This deviation-mitigating effect reduces (15) and further relaxes the ICC.

However, the third channel works against these two positive effects. The bracketed term in (14), \( \left[ (P^m - c_i) Q^m - (P^N - c_i) Q^N \right] \), decreases with mergers, because \( P^N \) increases and narrows its gap with \( P^m \). As market structure becomes closer to monopoly, the Nash profit eventually converges to the monopoly level, which obviates the need for collusion. The convergence of market structure decreases the gains from collusion (14) and makes the ICC more stringent. This convergence effect would not dominate the first two positive effects (of increasing \( s_i \)) under symmetric Cournot, but it could under cost asymmetry.46

46For example, suppose two low-cost firms (call them L1 and L2) and a high-cost firm (H) exist. \( P^N \) is
Fourth, mergers alter not only the number of firms but also their cost profile. Changing degrees of cost asymmetry affect $s_i$, which in turn affects the ICC through the three channels in the above. Moreover, if firm $i$ is merging, efficiency gains could directly affect the ICC by changing $c_i$ in (14) and (15).

Thus, the ICC could become either more stringent or less stringent as a result of mergers. Not only the magnitude of the coordinated effects but also their direction depends on market structure, cost asymmetry, and demand conditions. Hence, coordinated effects call for a careful quantitative analysis. Our explanation in the above highlights the intuition; see Appendix F.1 for a formal analysis.

7.2 The BASF-Takeda Merger

We start our empirical merger analysis with the specific case of BASF and Takeda, in which Takeda sold all of its vitamin businesses to BASF in 2001, that is, after the prosecution. Instead of 2001, we simulate a similar merger in 1990, a year before the vitamin C cartel started, so that we can study the impact of merger on collusive incentives during the period for which we have detailed evidence on the cartels.

Details are based on the UKCC’s (2001) 190-page report on the proposed acquisition:

- After the actual transaction in 2001, BASF shut down its own facilities and switched all production to Takeda’s plants. Hence, we assume the merged entity inherits Takeda’s cost structure, which was superior to BASF’s:

$$c_{\text{post basf}, \tau | t} = \min \left\{ c_{\text{pre takeda}, \tau | t}, c_{\text{pre basf}, \tau | t} \right\} = c_{\text{pre takeda}, \tau | t},$$

(16)

for all $\tau$ and $t$, where $c_{i, \tau | t}^{\text{pre}}$ and $c_{i, \tau | t}^{\text{post}}$ are pre- and post-merger marginal costs, respectively. Other, non-merging firms’ costs remain unchanged.

- We assume the existence of ample production capacity, because neither the long-run physical capacities nor the reported “nameplate” capacities were binding at any point during the sample period (see section 3.1).

\footnote{low when L1 and L2 are independent and compete against each other. $P^N$ will increase substantially if they merge, because the combination of L1 and L2 creates a near-monopoly player. Hence, the merged entity would not benefit as much from collusion as L1 or L2 independently would.}

\footnote{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. See Yamaguchi (2000). We thank Hidemaru Yamaguchi, health care and pharmaceutical research analyst at Citigroup Global Markets Japan, for sharing his knowledge on these transactions and industry history.}
The cartel in reality allocated quotas based on the historical market shares before the cartel (i.e., competitive outcomes in 1990). Accordingly, we construct a counterfactual version of “historical shares” and cartel quotas, by re-computing static Nash outcomes in 1990 between three firms (Roche, E. Merck, and the consolidated BASF-Takeda) instead of four in reality.

Under this setup, we recalculate streams of profits under the cartel, unilateral deviation, and static Nash, \( \left( \pi^c_{i,t}, \pi^d_{i,t}, \pi^N_{i,t} \right) \), for each \( i \in I, t \geq 1991, \) and \( \tau \geq t \). The repeated-games framework of section 5 allows us to convert these profits into the estimates of counterfactual ICCs. Assuming the cartel collapses when the collective ICC is violated for the first time, we calculate the frequency of its survival up to month \( t \) based on the 1,000 block-bootstrap ICC estimates,

\[
\text{Survival}_t = \prod_{m=\text{Jan-1991}}^t \Pr (ICC_m > 0) ,
\]

where \( ICC_m \) is the LHS of (13) in month \( m \). Our purpose is to construct an index that is informative about the cartel’s stability. This index summarizes the slackness of the ICC up to the current period, and provides a measure of how likely the ICC has been violated at least once, given the statistical uncertainty due to sampling error.

Figure 10 shows the cartel’s survival is more likely under the hypothetical merger in 1990—the dashed line labeled “merger (\( \sigma = 0 \))” at the top—than in the actual history without merger—the solid line labeled “Baseline (no merger).” However, its steep decline in 1996 suggests the coordinated effect would not have been sufficient to prevent the breakdown. This nuanced outcome turns out to be a norm rather than an exception, as the next two subsections show.

Readers who are familiar with Salant, Switzer, and Reynolds’s (1983) theoretical finding (that symmetric Cournot competitors do not have static incentive to merge) might question whether our quantity-game setup is suitable for merger simulations. We believe so for two reasons. First, their “no incentive to merge” results rely on symmetry and do not extend to settings with asymmetric firms. Perry and Porter (1985) showed this limitation in a Cournot game with asymmetric costs, and Deneckere and Davidson (1986) did so for a differentiated-good Bertrand case. Our case involves firms with asymmetric costs as well, and our calculations confirmed BASF and Takeda would have (static) incentives to merge (i.e., BASF-Takeda’s combined Cournot profits in 1991 would have increased by 6% under merger). Second, BASF’s closure of its own facilities after the actual merger is indicative of excess capacity. BASF-Takeda’s combined outputs decreased after merger. These developments closely match the “rationalization of assets” in Farrell and Shapiro’s (1990) theoretical analysis of mergers in a Cournot model.

48 Readers who are familiar with Salant, Switzer, and Reynolds’s (1983) theoretical finding (that symmetric Cournot competitors do not have static incentive to merge) might question whether our quantity-game setup is suitable for merger simulations. We believe so for two reasons. First, their “no incentive to merge” results rely on symmetry and do not extend to settings with asymmetric firms. Perry and Porter (1985) showed this limitation in a Cournot game with asymmetric costs, and Deneckere and Davidson (1986) did so for a differentiated-good Bertrand case. Our case involves firms with asymmetric costs as well, and our calculations confirmed BASF and Takeda would have (static) incentives to merge (i.e., BASF-Takeda’s combined Cournot profits in 1991 would have increased by 6% under merger). Second, BASF’s closure of its own facilities after the actual merger is indicative of excess capacity. BASF-Takeda’s combined outputs decreased after merger. These developments closely match the “rationalization of assets” in Farrell and Shapiro’s (1990) theoretical analysis of mergers in a Cournot model.
Note: Each line plots the frequency of cartel survival under the counterfactual BASF-Takeda merger, with a specific degree of efficiency gain ("synergy"), based on the 1,000 block-bootstrap estimates of the model under $\beta = 0.85$. Appendix F.2 shows their underlying ICC estimates.

### 7.3 The Role of Efficiency Gains

Could efficiency gains help? Let $\sigma \in [0, 1)$ parameterize synergy, and modify the merged entity’s cost as follows:

$$c_{basf, t}^{\text{post}} = (1 - \sigma) \times \min \left\{ c_{takeda, t}^{\text{pre}}, c_{basf, t}^{\text{pre}} \right\} = (1 - \sigma) \times c_{takeda, t}^{\text{pre}}.$$  \hspace{1cm} (18)

In Figure 10, four other lines (labeled $\sigma = 0.05$–0.20) reflect such synergistic scenarios. The 5%-efficiency-gain ($\sigma = 0.05$) scenario performs the best in terms of the eventual survival frequency. However, it does not strictly dominate the no-merger benchmark, because the survival frequency between 1991 and 1994 is lower. Higher levels of synergy ($\sigma = 0.10, 0.15, 0.20$) turn out to destabilize the cartel from the beginning. Thus, synergies could hinder collusion.

Figure 11 compares individual ICCs across these scenarios with different levels of $\sigma$. The first set of bar graphs ("Actual") shows the four firms’ actual ICCs in August 1995 without merger. The second set of bars labeled “Merger ($\sigma = 0$)” represents the three firms’ ICCs in the no-synergy merger counterfactual in the previous subsection. The remaining sets of bars reflect counterfactuals with synergistic mergers (5%–50% cost reductions relative to Takeda’s pre-merger level).
Figure 11: Coordinated Effects at Different Levels of Synergy

Note: The bars indicate the actual and counterfactual ICC estimates as of August 1995, under $\beta = 0.85$. Figure 25 in Appendix F.2 shows the full time profile of collective ICC and confidence intervals.

The impact of $\sigma$ is highly asymmetric across firms. Roche’s ICC relaxes at higher $\sigma$. By contrast, E. Merck’s ICC tightens with $\sigma$. The main source of asymmetry is their competitive positions relative to BASF-Takeda. As $\sigma$ increases, the cost competitiveness of BASF-Takeda becomes more similar to that of Roche (a low-cost rival) but less similar to that of E. Merck (a high-cost rival). Hence, their synergistic merger would facilitate collusion with Roche, but not with E. Merck, and destabilizes the cartel overall.\footnote{Another interesting feature of the firm-level plot is that the ICC of BASF-Takeda changes non-monotonically in $\sigma$, with a peak at around $\sigma = 0.3$. As $\sigma$ increases from 0, its cost competitiveness approaches Roche’s low cost, which makes its incentives more closely aligned with Roche’s. BASF-Takeda’s cost matches Roche’s at around $\sigma = 0.3$, above which it would effectively overtake Roche’s lowest-cost position in the industry. Stronger BASF-Takeda would become less fearful about the prospect of breaking the agreement and competing against Roche, and increasingly disgruntled with the Roche-led scheme.}

These firm-level asymmetries make the impact of a merger on collusion non-monotonic with respect to efficiency gains.

7.4 Concentration, Asymmetry, and Coordinated Effects

The critical role of cost asymmetry becomes clearer when we consider other hypothetical mergers. Table 3 lists the six cases we simulate and their key statistics: (i) the post-merger number of firms, (ii) the post-merger Herfindahl-Hirschman Index (HHI), (iii) the cartel’s
Table 3: Mergers, Asymmetry, and the Incentive to Collude

<table>
<thead>
<tr>
<th>Merger scenario</th>
<th>Continuing firms</th>
<th>Number of firms*</th>
<th>HHI* ($ thousand)</th>
<th>ICC**</th>
<th>Frequency of collusion**</th>
</tr>
</thead>
<tbody>
<tr>
<td>No merger</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>4</td>
</tr>
<tr>
<td>Merger #1</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
<td>3</td>
</tr>
<tr>
<td>Merger #2</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
<td>3</td>
</tr>
<tr>
<td>Merger #3</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
<td>3</td>
</tr>
<tr>
<td>Merger #4</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>2</td>
</tr>
<tr>
<td>Merger #5</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
<td>2</td>
</tr>
<tr>
<td>Merger #6</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td>2</td>
</tr>
</tbody>
</table>

* As of December 1990 (i.e., immediately before the beginning of the vitamin C cartel).
** As of August 1995 (i.e., the vitamin C cartel’s final month of operation on record).

For example, Merger #1 corresponds to the BASF-Takeda merger (without synergy) in the previous subsections, where Roche, Takeda, and E. Merck effectively become the continuing firms. The static Nash equilibrium would have led to the HHI of 3,542 as of December 1990 (i.e., immediately before the cartel’s beginning), which is higher than 3,182 in the no-merger case. The collective ICC in August 1995 (i.e., the cartel’s final month of operation in reality) is $301,000 in our point estimate with \( \beta = 0.85 \). The frequency of collusion (i.e., the frequency that the contemporaneous ICC is not violated) is 90% based on the 1,000 block-bootstrap estimates.

Similarly, Mergers #2 and #3 lead to triopoly market structures in which E. Merck and Takeda are absent, respectively (we do not consider eliminating Roche, because it is the industry leader with superior competitiveness). The post-merger HHI of 3,675 and 3,913 suggest these mergers would have resulted in a higher concentration of market shares. However, the ICC estimates ($231,000 and $12,000) suggest the cartel would have become less stable. Thus, a smaller number of firms and higher HHI, which are usually associated with a less competitive market structure, could also simultaneously be associated with a less collusive incentive structure.

Mergers #4, #5, and #6 further illustrate the nuanced relationship between market structure and the incentive to collude. Each of these cases considers transactions that result in a duopoly, and hence the post-merger HHI (4,380–4,990) is generally higher than under Mergers #1, #2, and #3 (3,542–3,913). However, these concentrated market shares do not necessarily translate into cartel stability. The ICC point estimates under Mergers #5 and #6 are lower than in most of the previous examples, and the probability of \( ICC_{Aug-1995} > 0 \) is 0% in both cases. Merger #4 is the only case in which cartel stability improves, because
Figure 12: Cartel Stability under Six Different Mergers

Note: Each line plots the frequency of cartel survival under a specific merger scenario, based on the 1,000 block-bootstrap estimates of the model under $\beta = 0.85$. Appendix F.2 shows their underlying ICC estimates.

it creates a highly symmetric duopoly of Roche and Takeda.\textsuperscript{50}

Figure 12 summarizes the time profile of cartel stability across these scenarios. Mergers #1 and #4 strictly outperform the no-merger benchmark, because they create market structures that are both more concentrated and \textit{more symmetric}. By contrast, the survival frequency under Merger #3 is relatively high until 1995 but quickly decreases to zero by 1996, because Roche’s ICC becomes stringent when it has to refrain from competing against high-cost rivals under deteriorating market conditions. Finally, the cost asymmetries under Mergers #5 and #6 create incentive problems from the beginning of the cartel.

Thus, although concentrated market structures may increase market power in terms of unilateral effects, both the sign and the magnitude of coordinated effects depend on the cost profile of all firms. These nuanced relationships between coordinated effects, concentration, and cost asymmetry highlight the importance of quantitative analysis.

\textsuperscript{50}Its ICC point estimate is not the highest of all cases, but its variability is considerably smaller (see Appendix F.2). Hence, none of our 1,000 simulations violated the ICC as of August 1995.
8 Conclusion

This paper shows a simple repeated-games model 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) allow us to conduct prospective merger simulations concerning coordinated effects. Such quantification is an important step to complement the unilateral-effect analysis and the theoretical “checklist” of facilitating factors.

In the course of empirical analysis, we also effectively “tested” one of the most fundamental predictions of repeated games–that cooperation breaks down when the incentive compatibility is violated. Repeated-games theory is known for its conceptual sophistication but is considered difficult for empirical implementation. With suitable data and institutional details, however, the framework is useful for issues of practical importance.

We focused on the vitamins case because of its historical importance and its relatively simple institutional setting. We expect this paper’s approach will be useful for studying other markets with similar characteristics, including many industrial markets for relatively homogeneous products. These markets have been prominent on the list of cartel investigations by the antitrust authorities and hence appear immediately relevant.\(^{51}\)

\(^{51}\)For example, see http://ec.europa.eu/competition/cartels/cases/cases.html.
# Appendix (For Online Publication)

This Appendix contains supplementary materials for sections 2 through 7.

## Table 4: Contents of the Appendix

<table>
<thead>
<tr>
<th>Appendix</th>
<th>Contents</th>
<th>Corresponding main text</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>Institutional and technological background</td>
<td>Section 2</td>
<td>A-1</td>
</tr>
<tr>
<td>A.1</td>
<td>Factual background on the cartels</td>
<td></td>
<td></td>
</tr>
<tr>
<td>A.2</td>
<td>Economic liberalization and SOEs in China</td>
<td></td>
<td></td>
</tr>
<tr>
<td>A.3</td>
<td>Vitamin production technology</td>
<td></td>
<td></td>
</tr>
<tr>
<td>B</td>
<td>Data considerations</td>
<td>Section 3</td>
<td>A-10</td>
</tr>
<tr>
<td>B.1</td>
<td>Prices</td>
<td></td>
<td></td>
</tr>
<tr>
<td>B.2</td>
<td>Costs</td>
<td></td>
<td></td>
</tr>
<tr>
<td>B.3</td>
<td>Capacities</td>
<td></td>
<td></td>
</tr>
<tr>
<td>B.4</td>
<td>Within-cartel market shares</td>
<td></td>
<td></td>
</tr>
<tr>
<td>C</td>
<td>Empirical analysis (additional results)</td>
<td>Sections 4 and 5</td>
<td>A-13</td>
</tr>
<tr>
<td>C.1</td>
<td>“Upward-sloping” demand curve?</td>
<td></td>
<td></td>
</tr>
<tr>
<td>C.2</td>
<td>The EM algorithm</td>
<td></td>
<td></td>
</tr>
<tr>
<td>C.3</td>
<td>Non-Roche firms’ marginal-cost estimates</td>
<td></td>
<td></td>
</tr>
<tr>
<td>C.4</td>
<td>ICC estimates under different levels of discounting</td>
<td></td>
<td></td>
</tr>
<tr>
<td>C.5</td>
<td>Demand and costs (other vitamins)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>D</td>
<td>Key modeling assumptions</td>
<td>Section 6.1</td>
<td>A-16</td>
</tr>
<tr>
<td>D.1</td>
<td>Quantity game</td>
<td></td>
<td></td>
</tr>
<tr>
<td>D.2</td>
<td>Delayed perfect monitoring</td>
<td></td>
<td></td>
</tr>
<tr>
<td>D.3</td>
<td>Fringe supply</td>
<td></td>
<td></td>
</tr>
<tr>
<td>D.4</td>
<td>Nash reversion</td>
<td></td>
<td></td>
</tr>
<tr>
<td>E</td>
<td>Sensitivity analysis</td>
<td>Section 6.2</td>
<td>A-20</td>
</tr>
<tr>
<td>E.1</td>
<td>Endogenous fringe</td>
<td></td>
<td></td>
</tr>
<tr>
<td>E.2</td>
<td>Renegotiation of quotas</td>
<td></td>
<td></td>
</tr>
<tr>
<td>E.3</td>
<td>Cartel price below monopoly level</td>
<td></td>
<td></td>
</tr>
<tr>
<td>E.4</td>
<td>Monitoring lag</td>
<td></td>
<td></td>
</tr>
<tr>
<td>E.5</td>
<td>Rational, static, and adaptive expectations</td>
<td></td>
<td></td>
</tr>
<tr>
<td>E.6</td>
<td>Log-linear demand</td>
<td></td>
<td></td>
</tr>
<tr>
<td>E.7</td>
<td>Long-term contracts</td>
<td></td>
<td></td>
</tr>
<tr>
<td>F</td>
<td>Mergers and coordinated effects</td>
<td>Section 7</td>
<td>A-31</td>
</tr>
<tr>
<td>F.1</td>
<td>Formal analysis of mergers’ impact on the ICC</td>
<td></td>
<td></td>
</tr>
<tr>
<td>F.2</td>
<td>Counterfactual ICCs</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

## Appendix A: Institutional and Technological Background

### A.1 Factual Background on the Cartels

(a) The Nature and Reliability of the Evidence

The EC decided to impose fines on the vitamin cartel members on November 21, 2001, and published an 89-page report in 2003. Its section 1.5, entitled “The nature and reliability of the evidence” (recitals 435–546), assesses the quality of the evidence is high: (i) Direct, documentary evidence is ample; (ii) statements from at least 13 manufacturers are available; and (iii) these statements are coherent and corroborated by the direct evidence.
(b) Competition before and after the Cartels

Section 1.4.1.1 of the EC report, entitled “The origin and basic scheme of the cartels,” describes the competitive environment in the markets of vitamins A and E before the beginning of the cartels (recital 160). The corresponding sections for the other cartels provide similar background information: vitamin C (recital 424) and vitamin H (recital 484). Collusion had allegedly been attempted in the mid-1980s, but these efforts did not seem effective or sophisticated. A steady drop in prices during the 1980s and the weakness of the dollar in 1989 and 1990 led to low profitability for Roche (e.g., recitals 293–295 on vitamin B5).

EC’s (2003) descriptions are sparse regarding the competitive environment after the cartels. Nevertheless, the precipitous drops in prices after the six short-lived cartels led to the post-cartel price levels that were 20%–50% lower than their respective pre-cartel levels. The UKCC’s (2001) merger report studies vitamins B2 and C in depth, and confirms “very active price competition” for both products since 1995 (p. 69, paragraph 4.94).

(c) Volume-Control Mechanism: “Budgets”

Bulk vitamins are sold in different forms, so the figures were expressed in terms of “pure” vitamins. Each producer converted its sales of diluted product into the equivalent of 100% product (EC 2003, recital 488). The “Basic scheme” sections of the EC report (sections 1.4.1.1 and 1.4.7.2) summarize the “budget” systems: vitamins A and E (recitals 164 and 170–171) and vitamin C (recitals 392–398). The following descriptions are particularly informative for our modeling purposes:

- “The accepted principle on which the cartel in vitamin C was based was that the existing worldwide market share of the four producers should be stabilised” (recital 392);
- “To set the quotas themselves, the participants first determined the total market on the basis of their sales and estimated sales by the Chinese and East Europeans producers of vitamin C. Expected sales by third parties were deducted, the remainder of the market being defined as the ‘available market’. Volume targets for each producer for the next period were set on the basis of their estimate of the ‘available market’” (recital 393);
- “The shares of the available market in 1990 (Roche 52 %, T 30 %, Merck 10 % and BASF 8 %) formed the basis for the allocations” (recital 394); and
- “There was to be ‘parallel development of sales and market share’, i.e. quotas were adjusted in volume terms to take account of increased demand while maintaining the same percentage shares and targets set each year by region. Sales would be monitored and the necessary corrections made quarterly” (recital 395).

(d) Meetings

Each cartel followed a slightly different protocol: (i) In the cartels of vitamins A and E, the European regional managers also had weekly and monthly telephone contact to exchange information on pricing and sales volumes, respectively; (ii) firms in the B1, B2, B5, B6, and B9 cartels met only bilaterally in most cases; (iii) the meetings took place only twice a year in the B6, D3, and H cartels; and (iv) Takeda negotiated with Roche in the vitamin B9 cartel as a representative of the
three Japanese firms in the Yōsankai group, which was a legacy of a government-led cartel in the past.

(e) Price Increases

Prices for vitamins A and E increased substantially between 1991 and 1994; the goal after 1994 was to maintain the achieved price levels. The vitamin C price increased as well, but in early 1993, Chinese products began to make incursions into the world vitamin C market, which led Roche to consider reducing the target price by 12%. They used rotating price-leadership in which “one producer should first ‘announce’ the increase” and then “the others would generally follow suit.” Recitals 200–204, 207, 210, and 221 in EC (2003) describe the manner in which the cartels increased prices of vitamins A and E in the first few years. Recitals 399–400 and 432–434 describe the price increases for vitamin C.

(f) Monitoring

The use of government statistics and “confirmation of sales” appear in the following recitals in EC (2003): vitamin B2 (289) and vitamin B9 (385). In the vitamin C cartel,

- “The four producers met again on 5 August 1993 in BASF’s offices in Frankfurt. Takeda has provided a detailed contemporaneous memorandum. Following an exchange of data it was confirmed that the 2.5 % cut referred to as a ‘voluntary target’ had more or less been achieved during the first six months of 1993. Price increases to DEM 25.00 were being implemented in Europe” (recital 439).

Firms could partially observe each other’s transaction prices with clients, albeit in a noisy manner, according to the following recitals from EC (2003): vitamin B6 (343), vitamin B9 (374), vitamin C (424, 425, and 435), and vitamin H (497 and 500). The UKCC’s (2001) merger report describes the nature of “publicly available” information about industry-wide pricing levels:

- “Although prices are not transparent, the parties told us that customers would contact several suppliers for quotations, and would quote Chinese (and other) prices in order to attain better deals from their suppliers. Consequently there was some knowledge of general prices within the industry. We were also made aware of two web sites, feedinfo.com and vetsquare.com, where a guide to current vitamin prices for feed-grade material is posted. Aventis told us that the sites obtain information by speaking to vitamins purchasers and finding the prices they have paid. As such, it is a publicly available guide to industry pricing” (4.89); and

- “Customers often source vitamins from more than one suppliers, and BASF argued that it was easy to switch between suppliers. (…) Customers confirmed they would shop around between suppliers” (4.90).

(g) Worldwide Markets and Production Locations

The EC report determines the relevant geographic markets are worldwide in its section 1.2.3, entitled “The relevant geographic market for vitamins A, E, B1, B2, B5, B6, C, D3, H, folic acid, beta-carotene and carotinoids” (recitals 69–73). The UKCC (2001) reached the same conclusion

A-3
(recitals 4.46 and 5.43). The worldwide nature of geographic markets is further reinforced by the presence of "brokers" or third-party traders of vitamins who engaged in cross-border arbitrage. The cartels were carefully calibrating "target prices" across countries in order to contain price differentials within 10%. See EC (2003), recitals 207, 223, and 321.

EC's (2003) section 1.2.4, entitled "Inter state trade," summarizes the plant locations in Europe and the prevalence of cross-border transactions (recitals 74–75). Although the EC report is mostly silent about production outside Europe, Japanese firms operated plants in (and exported from) Japan, according to UKCC (2001) and Bernheim (2002a).

(h) (Lack of) Interaction between Different Cartels

Theoretically, an agreement across multiple vitamins would seem possible. Roche organized 14 cartels, with BASF as its main partner (12 cartels), followed by Takeda (5 cartels), RP (3 cartels), and Daiichi (2 cartels), as Table 5 shows. However, the EC evidence does not suggest the existence of any contingency plans connecting multiple different cartels. Each cartel had its own protocol, operated separately, and ended at different times between 1994 and 1999. For example, Takeda was a member of five cartels (vitamins B1, B2, B6, B9, and C), but each had its own arrangement, and Takeda joined them on different occasions (recitals 252, 271, and 330). Another example is Daiichi, a key player in the vitamins B5 and B6 cartels. Chinese producers rapidly increased their vitamin B6 exports, and Roche and Takeda found Daiichi was "cheating" to match the Chinese fringe in 1993. The B6 cartel ended in 1994, but the cooperation between Daiichi and Roche continued in the B5 market until 1999 (recitals 343 and 347–350). This course of events is inconsistent with the idea of multi-market punishment.

Table 5: Global Market Shares (%) by Category in Early 1990s

| Firm    | Market: | A | B1 | B2 | B3 | B4 | B5 | B6 | B9 | B12 | C  | D3 | E  | H  | Caro- | All |
|---------|---------|---|----|----|----|----|----|----|----|-----|----|----|----|----| tinoi-|-----|
| 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   |    |


The only multi-product consideration that was documented in the EC report involved vitamin D3, which Roche regarded as part of vitamin A sales policy. The vitamin D3 cartel was unique in that the "agreement" seemed volatile and lacked a formal mechanism, unlike the other cartels (recitals 459, 490, and 497).
(i) Adjustments

The European regional meetings for vitamins A and E (Roche, BASF, and RP) operated a “compensation” scheme: “If at the end of the year a producer was substantially ahead of its quota, it had to purchase vitamins from the others in order to compensate them for the corresponding shortfall in their allocation” (recitals 196 and 225). However, their actual implementations were documented only for vitamin E.

Similar ideas were initially mentioned in the short-lived cartels for vitamins B9 and C as well (recitals 359, 367, and 398), but their constant quibbling about over-achievements casts doubt on whether they actually implemented compensations. For example, Takeda was constantly expressing its volume ambition in the vitamin C market, which irritated the three other members of the vitamin C cartel (recitals 429-430, 436, 441–442, and 444). The possibility of smooth year-end adjustments would seem remote in this market.

No such scheme seemed to exist for vitamins B5 and H (recitals 310 and 490).

(j) Six “Natural Deaths”

The typical reasons for the end of the six short-lived cartels include (i) the rapid growth of fringe supply, (ii) the slowdown of demand growth, and (iii) the allegations of cheating and other disagreements among the members, according to the following recitals from EC (2003): vitamin B1 (257 and 259), vitamin B2 (287 and 290–291), vitamin B6 (348–349), vitamin B9 (381–382), vitamin C (451 and 457), and vitamin H (510 and 513–515). Chinese producers, mostly state-owned enterprises (SOEs), significantly increased market shares in vitamins B1, B6, B9, and C. Il Sung, a Korean maker, was the fringe producer of vitamin H. 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.

(k) Ten “Antitrust Deaths”

According to Connor (2008, pp. 320–322), 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. See also EC (2003, recitals 90–97, 149–154, 227–233, and 328).

(l) Repeated Failures to Predict the Growth of Fringe

The following recitals from EC (2003) show the cartels had troubles predicting the growth rate of fringe supply and needed to update their forecasts in a staggered manner, even though their forecast of the worldwide demand seemed accurate: vitamin B1 (257–259 and 264), vitamin B2 (287), vitamin B6 (347–348), vitamin B9 (379–382), and vitamin C (430–431 and 454).
(m) Health Benefits of Vitamins

Dr. Robert Speights, a vitamin industry expert, prepared “Appendix E: Background on Vitamin Markets” for Bernheim’s (2002a) report. He states, “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” (Bernheim 2002a, Appendix A, pp. 9–10). By contrast, animal nutritionists conduct sophisticated cost-benefit analyses to determine the optimal mix of vitamins for each species.

(n) Substitutability between Vitamins

“Each group of vitamins (...) has specific metabolic functions and is therefore not interchangeable with the other groups” (EC 2003, recital 9). Likewise, each vitamin requires a specialized facility and production processes. Appendix A.3 lists raw materials and intermediate inputs for each vitamin. UKCC (2001) collected (and agreed with) the following statements from the industry:

- “BASF said that each vitamin fulfilled specific biochemical functions. As such, there was generally no demand-side substitutability for individual vitamins” (5.21);
- “BASF said that production plant was specific to each vitamin, and that there was no supply-side substitutability between vitamins. It followed that each vitamin was an individual product market” (5.22); and
- “BASF submitted that generally there was neither demand-side nor supply-side substitutability for individual vitamins. Each vitamin was used for its specific biochemical functions (see paragraph 5.21) and production of each vitamin required its own specific plant (see paragraph 5.22). Accordingly, at least each of the vitamins formed its own product market” (5.117).

(o) Bulk Vitamins Are Commodities

All sources including UKCC (2001), Bernheim (2002a), and EC (2003) characterize bulk vitamins as commodities for which price is the most important determinant of buyers’ purchasing decisions. The UKCC (2001) collects opinions from various parties, some of which claimed their own products’ superiority, but conclude the scope of product/brand differentiation is limited (paragraphs 4.89–4.90, 4.92, 4.94, and 5.128). Indeed, discussions at the cartel meetings suggest the buyers responded to prices as the most important factor (see Appendix A.1 (f)). Bernheim (2002a) compares vitamin prices across firms and found they were highly correlated and close to each other in levels. Appendix B.1 shows firm-level prices of four specific vitamin products.

A.2 Economic Liberalization and SOEs in China

Section 2.2 briefly described the policy change in China. Table 6 lists key historical 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, 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. A research team led by Professor Yin Guanglin of the Shanghai Research Center of Biotechnology of the Chinese Academy of Sciences
Table 6: Economic Liberalization and SOEs in China

<table>
<thead>
<tr>
<th>Year</th>
<th>Event</th>
</tr>
</thead>
<tbody>
<tr>
<td>1978</td>
<td>Deng Xiaoping starts economic liberalization policy</td>
</tr>
<tr>
<td>1980</td>
<td>Decentralization of budgetary control to local governments</td>
</tr>
<tr>
<td>1986</td>
<td>Legalization of private enterprises (with 8+ employees)</td>
</tr>
<tr>
<td>1991*</td>
<td>Soviet Union dissolved</td>
</tr>
<tr>
<td>1992*</td>
<td>Deng’s Southern Tour Speech endorsed private enterprises</td>
</tr>
<tr>
<td>1994*</td>
<td>Liberalization of prices and commerce complete</td>
</tr>
<tr>
<td></td>
<td>Modern corporate law and labor law became effective</td>
</tr>
<tr>
<td>1995*</td>
<td>SOEs started making losses collectively;</td>
</tr>
<tr>
<td></td>
<td>Restructuring and layoffs at SOEs permitted</td>
</tr>
<tr>
<td>1996</td>
<td>Last Five-Year Plan with production quantity targets</td>
</tr>
<tr>
<td>1997</td>
<td>Private sector endorsed as an “important part of socialist market economy”</td>
</tr>
<tr>
<td></td>
<td>Privatization of SOEs endorsed under the slogan “Grasp the large; let go of the small”</td>
</tr>
<tr>
<td>2002</td>
<td>Privatization of small and medium local SOEs (and layoffs at large SOEs) mostly complete</td>
</tr>
<tr>
<td></td>
<td>Accession to World Trade Organization</td>
</tr>
</tbody>
</table>

Note: * indicates key events that are related to the fringe expansion. See Chen, Igami, Sawada, and Xiao (2018).

was “the first in the world to develop and commercialize the production of vitamin C by a partial fermentation process” (UKCC 2001, p. 186). Roche “took a license over the 2 step process to keep it off the market outside of China” but “never used the technology” (p. 186).

Chinese entry into the export market showed staggered patterns for regulatory reasons. The US 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.

Four major firms appear on records in the 2000s (UKCC 2001, Bernheim 2008):

- North East Pharmaceutical is an SOE located in Shenyang, Liaoning Province, with its historical roots in Takeda’s pre-war plant.

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

- 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.

- 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.

The reported number of Chinese firms in the mid 1990s varies between 18 and 28, even within the same report by the UKCC (2001). Most of the private entrants had 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 Vitamin Production Technology

As section 2.3 briefly described, each vitamin requires highly specific inputs and production processes. Table 7 lists raw materials and intermediates for each vitamin. The existing technology is mature and common across firms within each category. Figure 13 shows the vitamin C production processes at Roche and Takeda. Only a minor difference exists between their procedures because both firms use the same underlying technology. By contrast, Chinese firms use a radically new approach, the two-step fermentation method.

Table 7: Key Chemical Ingredients Required for Vitamin Synthesis

<table>
<thead>
<tr>
<th>Vitamin</th>
<th>Intermediates</th>
<th>Raw Materials</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>Pseudoionone</td>
<td>Acetone, acetylene, isobutylene, butenediol, formaldehyde</td>
</tr>
<tr>
<td>B1</td>
<td>Grewe diamine</td>
<td>Ethylene, Prymidine, malononitrile, acrylonitrile, carbon monoxide, cetamidine, butyrolactone, methyl acetate, hydrochloric acid, ammonia, carbon disulphate</td>
</tr>
<tr>
<td>B2, synthetic</td>
<td>Ribose</td>
<td>Sugars</td>
</tr>
<tr>
<td>B2, fermented</td>
<td>Methylglutaronitrile, beta picoline, 3-cyanopyridine, methylethylpyridine</td>
<td>Ethylene, nitric acid, farmaldehyde, ammonia</td>
</tr>
<tr>
<td>B4</td>
<td>Trimethylamine</td>
<td>Hydrochloric acid, ethylene oxide</td>
</tr>
<tr>
<td>B5</td>
<td>Pantolactone, beta-alanine</td>
<td>Iso butryaldehyde, hydrogen yanoide, hydrochloric acid, acrylonitrile, ammonia, caustic soda, calcium hydroxide</td>
</tr>
<tr>
<td>B6</td>
<td>Pantolactone, beta-alanine</td>
<td>Oxazole, dienophile</td>
</tr>
<tr>
<td>B9</td>
<td>Pantolactone, beta-alanine</td>
<td>Acetone or acrolein, chlorine gas, guanadine, cyanoethyl acetate, sodium ethoxide, nitric acid, hydrogen gas, glutamic acid, benzoic acid</td>
</tr>
<tr>
<td>B12, fermented</td>
<td>Sorbitol</td>
<td>Sugars, nitrogen compounds</td>
</tr>
<tr>
<td>C</td>
<td>Glucose</td>
<td>Cholesterol</td>
</tr>
<tr>
<td>D3</td>
<td>Cholesterol</td>
<td>Acetone, acetylene, isobutylene, napha, formaldehyde</td>
</tr>
<tr>
<td>E, synthetic</td>
<td>Isophytol, trimethylhydroquinone</td>
<td>Furnaric acid or diketene, cysteine, thiophene, phosgene gas</td>
</tr>
<tr>
<td>H</td>
<td>Thiolactone</td>
<td>Acetone, ocetyene, triphenylphosphine</td>
</tr>
</tbody>
</table>

*Source: Bernheim (2002a), Connor (2007, 2008).*
Figure 13: Production Processes for Vitamin C

**Note:** DAS and MAS stand for Diacetone-L-Sorbose and Monoacetone-L-Sorbose, respectively.  
**Source:** Figures 6-2, 6-3, and 6-4 in Part IV (Vitamin C), Appendix E of Bernheim (2002a).
Appendix B: Data Considerations

B.1 Prices

Section 2.3 described the similarity in prices across firms and different grades. Figure 14 shows firm-level prices of four specific vitamin products, which are highly correlated and close to each other in levels. Prices are similar across firms because the underlying production technologies are the same, different “grades” are derived from the same base products (e.g., vitamin C in its pure crystal form), and the buyers aggressively shop around for lower prices due to the undifferentiated nature of bulk vitamins.

![Figure 14: Examples of Vitamin Prices by Firm and Grade](image)

Note: BASF’s vitamin C price might appear lower than its rivals', but EC’s (2003) documentary evidence suggests the appearance of discrepancy is probably not real. BASF’s internal accounting convention for this product seemed to generate constant and systematic discrepancies between its data and the other firms’ data in the order of 5%–10%, even though their actual transaction prices were virtually identical.

Any remaining differences are likely to reflect frictional factors such as the following. 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 confounding factors, the graphs exhibit limited discrepancies in the range of approximately 5% for over 15 years.

B.2 Costs

Bernheim (2002b) uses Roche’s internal cost data to investigate the relationship between the unit cost and the “nameplate” capacity utilization. Figure 15 plots Roche’s variable cost curves at production facilities in two different locations, Scotland (Dalry) and New Jersey (Belvidere). Three patterns emerge.

First, the nameplate capacity utilization could exceed 100%. 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), because of ad hoc definitions that vary across time and plants (see Appendix B.3).

Figure 15: Average Variable Cost and Reported Capacity Utilization by Plant

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

Second, the cost structures are remarkably similar across these plants. We have 16 annual observations for Dalry (their mean is $4.43 and standard deviation is $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 conventional significance levels.

The third noteworthy feature is the flat slopes of these two cost curves. Linear regressions do not reject the null hypothesis that the slope coefficients equal zero at conventional significance levels, either at the individual plant level or in the pooled sample. This finding motivates our modeling assumption (constant marginal cost) in section 4.2.52

---

52This observation does not necessarily contradict the conventional notion that “increasing returns to
B.3 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. Nevertheless, distinguishing between three different notions of capacity (two conceptual and one empirical) is useful for clarifying the considerations underlying our modeling choices.

First, “long-run” capacity (i.e., hard limits on output quantities defined by the physical sizes of production facilities) 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 (Bernheim 2002a, Appendix E).

Second, “short-run” capacity would correspond to Kreps and Scheinkman’s (1983) notion of quantity pre-commitment in our context. The closest empirical counterpart to this concept would be monthly production schedules, because smooth operations of chemical plants require orderly work shifts and timely procurement of raw materials.

Third, the reported nameplate capacities in the expert reports correspond to neither of these economic concepts, and would vaguely represent a noisy measure of medium-run production plans. 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 that “capacity is an opinion” whereas production is a fact.53 Thus, we follow Bernheim (2002a) in not using this ambiguous measure, and interpret the actual production as a reflection of the firms’ short-run production plans.

Table 8: Within-Cartel Market Shares

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Output (1,000 MT)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Roche</td>
<td>22,180</td>
<td>23,153</td>
<td>23,664</td>
<td>21,616</td>
<td>21,816</td>
<td>21,176</td>
<td>23,058</td>
<td>26,497</td>
<td>26,980</td>
</tr>
<tr>
<td>Takeda</td>
<td>12,815</td>
<td>13,086</td>
<td>14,199</td>
<td>13,209</td>
<td>13,089</td>
<td>12,573</td>
<td>11,529</td>
<td>11,854</td>
<td>12,070</td>
</tr>
<tr>
<td>E. Merck</td>
<td>4,436</td>
<td>4,033</td>
<td>4,141</td>
<td>4,203</td>
<td>4,363</td>
<td>3,970</td>
<td>4,069</td>
<td>4,881</td>
<td>4,970</td>
</tr>
<tr>
<td>BASF</td>
<td>3,450</td>
<td>3,523</td>
<td>2,958</td>
<td>3,603</td>
<td>3,117</td>
<td>3,309</td>
<td>2,713</td>
<td>3,486</td>
<td>4,260</td>
</tr>
<tr>
<td>Cartel total</td>
<td>42,881</td>
<td>44,795</td>
<td>44,962</td>
<td>42,631</td>
<td>42,385</td>
<td>41,028</td>
<td>41,370</td>
<td>46,719</td>
<td>48,280</td>
</tr>
</tbody>
</table>

Within-cartel share (%)

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Roche</td>
<td>51.7</td>
<td>51.7</td>
<td>52.6</td>
<td>50.7</td>
<td>51.5</td>
<td>51.6</td>
<td>55.7</td>
<td>56.7</td>
<td>55.9</td>
</tr>
<tr>
<td>Takeda</td>
<td>29.9</td>
<td>29.2</td>
<td>31.6</td>
<td>31.0</td>
<td>30.9</td>
<td>30.6</td>
<td>27.9</td>
<td>25.4</td>
<td>25.0</td>
</tr>
<tr>
<td>E. Merck</td>
<td>10.3</td>
<td>11.2</td>
<td>9.2</td>
<td>9.9</td>
<td>10.3</td>
<td>9.7</td>
<td>9.8</td>
<td>10.4</td>
<td>10.3</td>
</tr>
<tr>
<td>BASF</td>
<td>8.0</td>
<td>7.9</td>
<td>6.6</td>
<td>8.5</td>
<td>7.4</td>
<td>8.1</td>
<td>6.6</td>
<td>7.5</td>
<td>8.8</td>
</tr>
<tr>
<td>Cartel total</td>
<td>100.0</td>
<td>100.0</td>
<td>100.0</td>
<td>100.0</td>
<td>100.0</td>
<td>100.0</td>
<td>100.0</td>
<td>100.0</td>
<td>100.0</td>
</tr>
</tbody>
</table>

Note: The cartel period is January 1991–August 1995.

B.4 Within-Cartel Market Shares

Cartel quotas are based on market shares in the base year, which is 1990 in the case of vitamin C. Table 8 shows the member firms’ within-cartel shares are remarkably stable during the cartel period. The scale “characterize the production technologies for bulk chemicals. Firms do need to build large facilities (i.e., install some minimum-efficient physical capacity) to achieve these cost curves.

53 Transcript of McClymont’s deposition (p. 94), cit. in Landes, Sider, and Bamberger (2002, p. 49).
period between 1991 and 1995. In particular, Roche’s share stays within a 1% range of 51.7% in 1990. The other members also stick to their respective base-year shares, although their smaller scale of production seems to lead to slightly higher variability.

Appendix C: Empirical Analysis (Additional Results)

C.1 “Upward-Sloping” Demand Curve?

In section 4.1, we mentioned the problems for a standard IV regression in our data. Figure 16 (left) shows a scatter plot of $P_t$ and $Q_t$ exhibits an “upward-sloping” curve, presumably because of the secular growth of demand (i.e., the increase of demand shifter, $X_t$, over time). The only “downward-sloping” portion of data is the years surrounding the breakdown of the vitamin C cartel in August 1995, but this subsample is too small.

The right panel of Figure 16 shows a scatter plot of $P_t$ and $c_{i,t}$. Roche’s unit cost has a respectable fit with $R^2 > .5$, but its range is confined to approximately $2.00. This limited variability makes the estimation of a standard (downward-sloping) demand curve difficult in the presence of strong multicollinearity between $P_t$ and $X_t$.

C.2 The EM Algorithm

We estimate demand and costs by minimizing the sum of squared residuals in the four firms’ Cournot FOCs (equations 3 and 7), which requires monthly observations of $q_{i,t}$. Column (1) of Table 2 uses imputed monthly outputs based on an equal split of the annual data,

$$q_{i,t} = \frac{1}{12} q_{i,y(t)}; \quad (19)$$

The right panel of Figure 16 shows a scatter plot of $P_t$ and $c_{i,t}$. Roche’s unit cost has a respectable fit with $R^2 > .5$, but its range is confined to approximately $2.00. This limited variability makes the estimation of a standard (downward-sloping) demand curve difficult in the presence of strong multicollinearity between $P_t$ and $X_t$. 

Figure 16: Demand Growth and Limited Cost Variation

The right panel of Figure 16 shows a scatter plot of $P_t$ and $c_{i,t}$. Roche’s unit cost has a respectable fit with $R^2 > .5$, but its range is confined to approximately $2.00. This limited variability makes the estimation of a standard (downward-sloping) demand curve difficult in the presence of strong multicollinearity between $P_t$ and $X_t$. 

C.2 The EM Algorithm

We estimate demand and costs by minimizing the sum of squared residuals in the four firms’ Cournot FOCs (equations 3 and 7), which requires monthly observations of $q_{i,t}$. Column (1) of Table 2 uses imputed monthly outputs based on an equal split of the annual data,
where \( y(t) \) denotes the year to which month \( t \) belongs. This approach is simple but ad hoc and potentially introduces monthly measurement error into the RHS of the regression equation.

Column (2) uses an estimation procedure that addresses this problem more systematically by using the EM algorithm (c.f., Little and Rubin 2002, section 8.2). We set the initial values \( q_{i,t}^{(0)} \) for all \( i \) and \( t \) according to (19). In the first step (the M step), we estimate the parameters

\[
\theta^{(1)} = \arg \min_{\theta} \sum_{t} \sum_{i} \hat{\eta}_{i,t}^2 \left( q^{(0)}; \theta \right) \quad \text{if } I_t = 0,
\]

where \( \hat{\eta}_{i,t} \) is the residual from (3) and (7), and \( q^{(0)} \) collectively denotes \( q_{i,t}^{(0)} \) for all \( i \) and \( t \).\(^{54}\) In the second step (the E step), we find the conditional expectation of the missing monthly data \( q_{i,t} \) given \( \theta^{(1)} \) and the annual observations \( q_{i,y(t)} \):

\[
q^{(1)} = \arg \min_{q} \sum_{t} \sum_{i} \hat{\eta}_{i,t}^2 \left( q; \theta^{(1)} \right) \quad \text{if } I_t = 0,
\]

subject to \( \sum_{t \in y(t)} q_{i,t} = q_{i,y(t)} \) for all \( i \) and \( y(t) \), and subject to \( q_{i,t} \geq 0 \) for all \( i \) and \( t \).

Subsequently, we iteratively update \( \theta^{(2)} \), \( q^{(2)} \), \( \theta^{(3)} \), \( q^{(3)} \), and so on until convergence.

Figure 17: The EM Algorithm

Figure 17 (left) plots the estimate of the demand slope at each iteration, which suggests most of the improvement in fit occurs within the first few iterations. The right panel shows the imputed \( q_{i,t} \) (solid lines) is visibly different from its initial value (dashed lines), but most of the variation is across years, which explains why the two estimates in Table 2 are similar.

\(^{54}\)We assume the error term \( \eta_{i,t} \) (the measurement error in \( c_{i,t} \)) is normally distributed, in which case the log likelihood function is concentrated with respect to \( \theta \) and can be maximized by minimizing the sum of squared residuals (c.f., Hayashi 2000, section 1.5).
C.3 Non-Roche Firms’ Marginal-Cost Estimates

Non-Roche firms’ cost data in Bernheim (2002a) are incomplete. Consequently, we estimate their implied marginal costs in section 4.3. Figure 18 plots their marginal-cost estimates along with Takeda’s (incomplete) cost data. Among the three non-Roche firms, only Takeda’s “data” appear in the vitamin C section of Bernheim (2002a), with either radical fluctuations or missing values. The estimated cost heterogeneity between Roche and its rivals seems less pronounced than their market-share heterogeneity. For example, Roche’s average output is almost 100% larger than Takeda’s during the 1990s; Roche’s cost is on average 34% (or $3.36) lower than Takeda’s in the same period. We regard these cost differences as a reflection of logistics costs, local input-price conditions in each country, and other idiosyncratic components of productive efficiency that are specific to firms and plants (e.g., managerial practices).

Figure 18: Non-Roche Firms’ Marginal Costs

Note: Takeda’s cost “data” exhibit implausible fluctuations and are lacking for most of the 1980s. We plot them here only to illustrate the data problems regarding non-Roche firms’ costs.

C.4 ICC Estimates under Different Levels of Discounting

Figure 19 shows Roche’s ICC (i.e., the most stringent of the individual ICCs in 1995 and 1996) for a range of $\beta \in \{0.75, 0.80, 0.85, 0.90, 0.95\}$. Small numerical differences are visible, but the qualitative patterns hardly vary. Some readers might wonder whether $\beta \leq 0.95$ is a reasonable level of patience for the global chemicals/pharmaceutical sector. We believe so for two reasons. First, the cartel was organized by the division managers, not the CEOs or the institutional shareholders of these firms.\(^{55}\) Hence, the $\beta$ in our case would also reflect their personal time horizons (e.g., bonuses, promotions,

\(^{55}\)Recall that the heads of the vitamins divisions of Roche and BASF, as well as of RP’s Animal Nutrition division, managed the cartels (section 2.1).
and retirement). Second, $\beta$ may incorporate 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. Thus, the subjective $\beta$ of individual managers should be lower than that implied by the risk-free rates as in the macroeconomics literature (e.g., 0.95 or 0.99).\textsuperscript{56}

Figure 19: Estimates of Roche’s ICC under Different Levels of $\beta$

![Graph showing estimates of Roche’s ICC under different levels of $\beta$.]

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

C.5 Demand and Costs (Other Vitamins)

Table 9 reports the demand and cost estimates for vitamins A and E, and beta carotene. Different vitamin markets operate on different scales (see Figure 4 in section 3), which is why the estimates of $\frac{dP}{dQ}$ and $\gamma_i$ appear in various scales. Note the standard error of $\hat{\gamma}_{basf}$ is large for beta carotene, because BASF was scaling up the production of this product in the 1980s (i.e., $q_{basf,t}$ varied substantially across $t$).

Appendix D. Key Modeling Assumptions

We use a standard model of a quantity game with (delayed) perfect monitoring. Appendices D.1 and D.2 critically assess these assumptions, consider alternatives, and discuss why our main specification is still a useful approximation. Appendices D.3 and D.4 further elaborate on two other aspects of our model: fringe supply and Nash reversion.

\textsuperscript{56}The fact that some of the senior executives eventually went to American jail portrays the risky nature of the enterprise at personal levels.
Table 9: Estimates of Demand and Costs (Vitamins A and E, and Beta Carotene)

<table>
<thead>
<tr>
<th>Market</th>
<th>Vitamin A</th>
<th>Vitamin E</th>
<th>Beta Carotene</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(1)</td>
<td>(2)</td>
<td>(3)</td>
</tr>
<tr>
<td>$dP/dQ$</td>
<td>-0.076</td>
<td>-0.033</td>
<td>-47.060</td>
</tr>
<tr>
<td></td>
<td>(0.001)</td>
<td>(0.003)</td>
<td>(6.311)</td>
</tr>
<tr>
<td>$\gamma_{basf}$</td>
<td>11.910</td>
<td>8.526</td>
<td>232.544</td>
</tr>
<tr>
<td></td>
<td>(0.248)</td>
<td>(0.475)</td>
<td>(480.190)</td>
</tr>
<tr>
<td>$\gamma_{cp}$</td>
<td>16.869</td>
<td>11.014</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>(0.483)</td>
<td>(0.493)</td>
<td>(-)</td>
</tr>
<tr>
<td>$\gamma_{cisai}$</td>
<td>-</td>
<td>11.921</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>(-)</td>
<td>(0.576)</td>
<td>(-)</td>
</tr>
<tr>
<td>Number of observations</td>
<td>60</td>
<td>60</td>
<td>72</td>
</tr>
</tbody>
</table>

Note: Standard errors in parentheses are based on 1,000 block-bootstrap samples, where each block consists of 12 consecutive months of a calendar year. We use the EM algorithm to impute monthly outputs as part of the estimation procedure.

D.1 Quantity Game

We assume the stage game is a quantity game, which is a standard IO model for homogeneous-good industries such as bulk chemicals. A usual criticism of the Cournot model, dating back to the time of Bertrand, is that firms in the real world usually choose prices as well as quantities. We consider two alternatives: a sequential quantity-price (Q-P) game and a simultaneous Q-P game.

**Tension between Cournot and Kreps-Scheinkman**

Kreps and Scheinkman (1983, henceforth KS) propose to interpret the Cournot quantities as the subgame-perfect equilibrium quantities of their sequential Q-P game with the parallel rationing rule. Note the KS (sequential Q-P) model is not exactly the same as the Cournot model, in two aspects. One is the lack of strategic equivalence. They establish the equivalence of equilibrium outcomes between a quantity game and a sequential Q-P game, but not their equivalence in strategic form. This lack of strategic equivalence is not an immediate problem for the calculation of the static Nash profits ($\pi^N$). However, it becomes a problem for the computation of deviation profits, because $\pi^D$ in a quantity game might be different from $\pi^D$ in a sequential Q-P game. Another difference is that the KS model assumes the first-stage (pre-committed) quantities become public information in the second stage in which firms choose prices. This assumption may not be a bad one if physical capacities play the role of quantity pre-commitment (i.e., within a relatively long time interval), but our empirical context features relatively short-run “production plans” as quantity decisions. KS’s public-information assumption seems strong for an environment with lagged perfect monitoring.

**Two Alternatives to Quantity Game**

First, let us consider a sequential Q-P game (i.e., the KS model). Firms choose their respective “production plans” (i.e., pre-committed quantities) in the first stage and set their prices in the second stage. We investigate three different informational settings: (i) Each firm cannot observe its rivals’ production plans at all; (ii) the production plans are not common knowledge, but each firm has some incomplete information about them; and (iii) the production plans are common knowledge.

- The static Nash equilibrium of setting (i) induces a Bertrand outcome, which is inconsistent with the industry background of homogeneous goods and the existence of substantial markups in the data (see section 3.2).
The analysis of setting (ii) is empirically infeasible under most circumstances, because the analyst would need to know each firm’s beliefs about the production plans of the opponents, which is not recorded in usual datasets.

Setting (iii) creates an interesting problem: What if a firm notices that an opponent’s production plan does not follow the cartel agreement? One possibility is that, from the subsequent period, the firms play the KS equilibrium forever as a punishment phase. With this grim prospect for the future, the firms in the current period would play the pricing subgame (given the production plans) without any consideration about the continuation payoffs. This subgame corresponds to the KS subgame after some firm chooses a quantity that does not correspond to the on-path quantity of the subgame-perfect equilibrium. Solving for the outcome of this subgame requires solving a Bertrand game with heterogeneous costs and capacity constraints, which may give rise to multiple equilibria and make the problem intractable.

Second, let us consider a simultaneous Q-P game. The Nash equilibrium of this game induces a Bertrand outcome, which leads to the same empirical problem under setting (i) in the above. Thus, these alternative stage-game specifications have their own problems. They are either conceptually intractable, empirically intractable, or inconsistent with the institutional context of the vitamins markets. Consequently, we use the quantity game as our preferred specification.

Results in Favor of the Quantity Game At this point, we would like to highlight three empirical findings that suggest a quantity game is a useful approximation in our context. First, the implied marginal costs of non-Roche firms (for which the internal cost data are limited) under the Cournot assumption are reasonably close to each other and Roche’s data, which matches the industry context of mature production technologies (see Appendix C.3). Second, the implied monopoly prices based on our demand estimates (which in turn are based on the maintained assumption of a quantity game during the non-cartel periods) match closely with the actual price during the cartel period. These implied monopoly prices are not constrained to fit the cartel prices in the data and hence constitute an independent means to assess the validity of the Cournot assumption (see section 4.4). Third, our estimates suggest a range of discount-factor levels exists such that the ICC for the vitamin C cartel is violated around the time of its actual collapse, and that the ICCs for the other three vitamin cartels are always satisfied (see section 5.2 and Appendix C.4).

In summary, we regard a quantity game as a useful approximation to the spot-market transactions in the vitamins markets. We are aware of the Bertrand critique of the Cournot model, and of the fact that oligopolistic firms in the real world may choose prices as well as quantities. Alternative specifications such as Q-P games might be more “realistic,” but the lack of tractability makes them less amenable to empirical implementation.

D.2 Delayed Perfect Monitoring

Why Did Firms Wait? The second important assumption is that no punishment is conducted until non-compliance is verified with the government statistics. We assume \( \dot{X}_t \) and \( P_t \) are commonly observed at the beginning and end of period \( t \), respectively. Because \( P_t = \frac{dP}{dQ} \left( Q_{\text{cartel},t} + Q_{\text{fri},t} - \dot{X}_t \right) \), the reader may notice the theoretical possibility that each firm should be able to infer \( Q_{\text{cartel},t} \) from \( \left( P_t, \frac{dP}{dQ}, Q_{\text{fri},t}, \dot{X}_t \right) \) 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 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 (see section 2.1).

How should we reconcile these two observations? Theoretically, we can consider augmenting the inverse demand with noise, $P_t = \frac{dP}{dQ_t} \left( Q_{\text{cartel},t} + Q_{\text{fri},t} - \bar{X}_t \right) + \nu_t$, where $\nu_t$ is an i.i.d. noise with $\mathbb{E}[\nu_t] = 0$. If $\nu_t$ has a full support, individual firms cannot identify $Q_{\text{cartel},t}$ 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 the model to incorporate this mean-zero error, $\nu_t$, does not change our analysis.\footnote{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 organization of the vitamin cartels as reported in EC (2003).}

**Why Did Firms Self-Report?** Why did firms bother to exchange self-reported sales records when more reliable information would become available within a few months anyway? Modeling communication is beyond the scope of this paper, but one theoretical possibility is that such communication may facilitate collusion. Awaya and Krishna (2016) derive conditions under which cheap talk on past sales increases collusive profits, assuming sales data remain private forever. Spector (2015) investigates the role of cheap talk when the underlying information becomes publicly available with a delay. These papers study settings that are slightly different from ours, and hence we may not directly apply their arguments to interpret the role of communication in our context, but they suggest possible directions for future research.

Another important paper that investigates the issue of monitoring in detail is Harrington and Skrzypacz (2011), which puts forth a very different model of collusion with private monitoring. Their model is based on the lysine and citric acid cartels, but they state its relevance to vitamins and other cartels as well. In their setting, firms agree on price and sales quotas, exchange self-reported sales, and make adjustments when sales are misaligned with quotas. The authors identify three ICCs associated with (i) choosing the collusive price, (ii) truthfully reporting sales, and (iii) making adjustments. They show (ii) is the binding ICC. By contrast, our specification features perfect monitoring, reflecting the prominent role of third-party verification of sales via government statistics in the vitamins case. We have chosen not to explicitly model the underlying communication and the three ICCs, not because we disregard their importance but because systematic records are unavailable regarding the details of self-reports and adjustments.

### D.3 Fringe Supply

We assume the cartel held and updated static expectations about $Q_{\text{fri},t}$ because EC (2003) suggests the cartels acted on such forecasts (see section 2.2). Even if we suppose instead that the cartel firms had had near-perfect foresight about $Q_{\text{fri},t}$, the following model could still rationalize “static expectations.”

Suppose each player belongs to one of the two possible types, “sophisticated” and “naïve.” The sophisticated type rationally expects that once the Chinese SOEs start growing, they will eventually violate the ICC in (13). Meanwhile, the naïve type holds the static belief and has to update it with news every year. In the best equilibrium for the cartel, sophisticated members will optimally choose to pretend to be naïve 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.

The intuition is the same as in Kreps, Milgrom, Roberts, and Wilson (1982). 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 (13). Likewise, the naïve type in their context means the “commitment” type to cooperation; the naïve type in our model would be the commitment type that follows the cartel agreement until (13) is violated in its static belief. In both cases, the pretense of naiveté can extend the length of cooperation periods, whereas the cooperation 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 Chinese SOEs suddenly expand and 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. naïve) in the above.

D.4 Nash Reversion

We specify the punishment takes the form of an infinite repetition of the static Nash equilibrium, because the cartel members communicated their understanding of the agreement as such (see section 2.1). Theorists have proposed more severe forms of punishment. For example, Abreu (1986) constructs a “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.

Appendix E: Sensitivity Analysis

The repeated-games framework in section 5.1 can be customized in many ways. This Appendix investigates the consequences of alternative specifications.

E.1 Endogenous Fringe

Let us endogenize fringe supply by modeling it as a lagged supply curve. The two graphs in Figure 20 plot the total fringe output against the lagged price (left) and the current price (right), respectively, in the vitamin C market. The fringe output strongly correlates with the lagged price ($R^2 = 0.57$) but only modestly with the current price ($R^2 = 0.24$), which is consistent with the context of staggered entry by Chinese suppliers (see section 2.2 and Appendix A.2).

Suppose the lagged supply curve adequately describes the endogenous supply response from China. With time, the cartel firms should be able to learn and estimate this fringe-supply function.

---

58See also Kreps and Wilson (1982) and Milgrom and Roberts (1982).

59This 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.
Eventually, they might even come up with a “better” cartel output level to keep the Chinese exports low (i.e., a dynamic version of limit pricing).

To investigate this possibility, suppose the price is at its stationary level, and the cartel firms play a Markov perfect equilibrium (MPE). Note an MPE is the adequate non-collusive equilibrium concept (corresponding to a static Nash in our baseline model), because now we assume the competitive fringe reacts to past prices and the cartel firms react to it. Could the cartel firms possibly agree on a new cartel scheme under this situation, with the reversion to MPE as a punishment? Our answer is “no,” as we explain in the following.

Let us fix $X_t = x$, $\frac{dP}{dQ} = \psi$, and $c_{i,t} = c_i$ at the 1998 levels. The model is the same as in section 4 except that firms rationally expect that the fringe supply reacts to past $K$-period prices linearly:

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

where $a_{fri}$ and $b_{fri,k}$ are coefficients of the fringe-supply curve.\(^{60}\) The inverse demand function for the total cartel output, $Q = \sum_{i \in I} Q_{i,t}$, is

$$P = -\psi \times (x - Q_{fri}) + \psi \times Q, \text{ or } Q = x - Q_{fri} + \frac{P}{\psi}. \quad (22)$$

We first derive the optimal cartel supply, where the players jointly maximize the cartel profit. The state variable for calculating the optimal cartel profit is $(P_{-k})_{k=1}^{K}$, where $P_{-k}$ is the price $k$ periods before, because the environment is stationary. The value function $V((P_{-k})_{k=1}^{K})$ satisfies

\(^{60}\) We do not allow the Chinese firms to join the cartel, because the incumbent firms in Europe and Japan characterized them as aggressive and untamable, according to UKCC (2001).
that the cartel chooses \( Q \) given (22) to maximize

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

(23)

where \( P_{-l:k} = (P_{-l}, P_{-l-1}, \ldots, P_{-k}) \). Note that \( Q_{fri} \) depends on \( (P_{-k})_{k=1}^K \). The FOC is

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

where \( V_k \) is the derivative of the \( k \)th argument. By the envelope theorem, we have

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

and so

\[
V_k (P_{+k-1:K+k}, P_{-1:K+k}) = -(P_{+k} - c) b_{fri,k} + \beta V_{k+1} (P_{+k+1}, P_{-1:K+k+1}),
\]

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

\[
\beta V_1 (P, P_{-1:K+1}) = -\sum_{k=1}^K \beta^k (P_{+k} - c) b_{fri,k}.
\]

(24)

Given (24), the FOC can be written as

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

or

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

(25)

We prove a linear solution exists:

**Lemma 1** There exists a linear solution for (23): \( P = a_{cartel} + \sum_{k=1}^K b_{cartel,k} P_{-k} \).

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

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

\[
a_{cartel} + \sum_{k=1}^K b_{cartel,k} P_{-k}
\]

(26)

and verify it. Given the guess, repeatedly substituting (26), we have

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

A-22
with \( \sum_{m=1}^{0} (b_{\text{cartal},m})^{k-m} = 1 \). Putting them back into the FOC (25) and matching the coefficient gives us the linear solution. Hence, the guess is verified. \( \blacksquare \)

**MPE** Suppose no cartel exists. Then each firm would like to maximize its profit by changing \( q_i \), taking the other firms’ outputs 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 to produce \( Q_{-i} \) and \( Q_{fri} \), respectively.

**Lemma 2** There exists a linear MPE, in which each firm \( i \) produces

\[
q_i (P_{-1:K}) = a^i_{MPE} + \sum_{k=1}^{K} b_{MPE,k} P_k + \gamma^i_{MPE} c_i.
\]

Note 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

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

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

\[
x - Q_{fri} - Q_{-i} + \frac{2P - c_i}{\psi} + \beta V_1^{i,MPE} (P, P_{-1:K+1}) = 0.
\]

By the envelope theorem, we have

\[
V_k^{i,MPE} (P_{-1:K}) = (P - c_i) \left( -b_{fri,k} - \frac{dQ_{-i}}{dP_k} \right) + \beta V_{k+1}^{i,MPE} (P, P_{-1:K+1})
\]

\[
= (P - c_i) \left( -b_{fri,k} - (n-1) b_{MPE,k} \right) + \beta V_{k+1}^{i,MPE} (P, P_{-1:K+1})
\]

and so

\[
\beta V_1^{i,MPE} (P, P_{-1:K+1}) = - \sum_{k=1}^{K} \beta^k (P_k + c_i) \left( b_{fri,k} + (n-1) b_{MPE,k} \right).
\]

Hence, the FOC is equivalent to

\[
P = \frac{\left| \psi \right| (x - Q_{fri} - Q_{-i}) + c_i - \left| \psi \right| \sum_{k=1}^{K} \beta^k (P_k + c_i) \left( b_{fri,k} + (n-1) b_{MPE,k} \right)}{2}.
\]
Under this cartel agreement, the optimal non-compliance profit is calculated as follows. Given on-path price \( P_{t-k} = \bar{P} \) for each \( k \), we calculate the sequence of cartel quantities and prices, \( Q_t = \bar{Q}(\bar{P}, \ldots, \bar{P}) \), \( P_t = \bar{P}(\bar{P}, \ldots, \bar{P}) \), \( Q_{t+1} = \bar{Q}(\bar{P}, \ldots, \bar{P}, P_t) \), \( \bar{P}_{t+1} = \bar{P}(\bar{P}, \ldots, \bar{P}, P_t) \), \( Q_{t+2} = \bar{Q}(\bar{P}, \ldots, \bar{P}, P_t, P_{t+1}) \), \( \bar{P}_{t+2} = \bar{P}(\bar{P}, \ldots, \bar{P}, P_t, P_{t+1}) \), and so on. Suppose firm \( i \) obtains the market share \( s_i \). Then its cartel profit is \( (P_t - c_i) s_i Q_t, (P_{t+1} - c_i) s_i Q_{t+1}, \) and so on.

**Cartel Profit** 

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

**Optimal Non-compliance** 

Under this cartel agreement, 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 \( (1 - s_i) \bar{Q}(\{P_{t-k}\}_{k=1}^K) \) for the next \( L \) periods \( t = t', \ldots, t' + L - 1 \), and then will switch to the MPE \( q_i(\{P_{t-k}\}_{k=1}^K) \) from \( t' + L \), where \( L \) denotes the monitoring lag. For each past price \( P_{t-1:-k} \) (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 \( s_i \) share in the cartel agreement, we have

\[
P = -\psi(x - Q_{fri}) + \psi(s_{-i} Q_{cartel} + q_i)
\]

with \( s_{-i} = 1 - s_i \). In other words,

\[
q_i = (x - Q_{fri}) + \frac{P}{\psi} - s_{-i} Q_{cartel}.
\]

Because \( Q_{cartel} \) is known, selecting \( q_i \) is equivalent to selecting \( P \). Hence, the optimal non-compliance (given \( Q_{cartel} \)) is to maximize

\[
V_{i,L}^{i,s} (P_{-1:-k}) = \max_{P} (P - c_i) \left[ x - Q_{fri} + \frac{P}{\psi} - s_{-i} Q_{cartel} \right] + \beta V_{i,MPE}^{i,s} (P, P_{-1:-k+1}) .
\]
From (27), we know $V_{1}^{i,MPE}(P, P_{-1:K-1})$. Hence, the FOC becomes

$$
P = \frac{|\psi| (x - Q_{frti} - s_{-i}Q_{cartel}) + c_{i}}{2}
- \frac{|\psi| \sum_{k=1}^{K} \beta^{n_{f}} (P_{+k} - c_{i}) (b_{frti,k} + (n - 1) b_{MPE,k})}{2}.
$$

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

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

$$
V_{i,L-1}^{i} (P_{-1:-K}) = \max_{P} (P - c_{i}) \left[ (x - Q_{frti}) + \frac{P}{\psi} - s_{-i}Q_{cartel} \right] + \beta V_{i,L}^{i} (P, P_{-1:K-1}).
$$

Because we now know $V_{i,L}$, we can solve for the optimal $P$ as a function of $Q_{frti}$, $Q_{cartel}$, 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 no $i$ and $\tau$ exist such that

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

for the on-path sequence of cartel quantities and prices, $Q_{t} = \bar{Q}(\bar{P}, ..., \bar{P})$, $P_{t} = P(\bar{P}, ..., \bar{P})$, $Q_{t+1} = \bar{Q}(\bar{P}, ..., \bar{P}, \bar{P}_{t})$, $P_{t+1} = P(\bar{P}, ..., \bar{P}, \bar{P}_{t})$, $Q_{t+2} = \bar{Q}(\bar{P}, ..., \bar{P}, \bar{P}_{t}, \bar{P}_{t+1})$, $P_{t+2} = P(\bar{P}, ..., \bar{P}, \bar{P}_{t}, \bar{P}_{t+1})$.

To empirically implement the model of collusion (with endogenous fringe) in the above, we specify $K = 2$ and regress $Q_{frti,t}$ on $P_{t-2}$ and $P_{t-1}$ to estimate the fringe-supply function (see Figure 20, left). We investigate whether any share profile $(s_{i})_{i \in I}$ exists with which (31) holds. However, we find no feasible $(s_{i})_{i \in I}$ exists for each $\beta \leq .9$. In particular, Roche’s ICC requires $s_{roche} > .8$, whereas Takeda’s ICC requires $s_{takeda} > .2$. These two conditions are incompatible.

Compared to the myopically 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 breaks the cartel’s ICC.

**E.2 Renegotiating Quotas**

Could the vitamin C cartel have optimally renegotiated quotas and avoided its collapse in 1995? Increasing Roche’s quota by 10% or more could relax its ICC sufficiently, but this level of quota transfer would be impractical. Either E. Merck or BASF would have to give up more than half of its allotment. Moreover, EC (2003) suggests Takeda was constantly trying to negotiate an increase of its own quota (see section 2.1). Hence, such renegotiations would have been difficult. This conclusion is further reinforced by our finding in Appendix E.1 that no quota reallocation could have saved the cartel in the late 1990s (under the alternative model with endogenous fringe).
E.3 Cartel Price below Monopoly Level

Let us investigate the sensitivity of our ICC estimates with respect to the level of cartel pricing. Instead of assuming perfect coordination, suppose the cartel’s collective outputs exceeded the monopoly level by 5%–15%. In Figure 21, the bold line (100%) is the baseline estimate with $\beta = 0.85$, whereas the five others reflect greater output levels of the cartel (105%–115%). The ICC would generally become more binding if the cartel coordinated to achieve less-than-monopoly price levels. The only exception is a slight outperformance of 105% in 1995. Thus, the cartel did not have a reason to set a target price below the monopoly level.

These outcomes are reasonable. 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 $\Delta V \equiv V_C - V_D$ might increase with a small change from the monopoly price. However, when the price departs significantly from the monopoly level, its first-order (negative) effect on $\pi_C$ kicks in to drive down $V_C$ along with $\Delta V$.

![Figure 21: Collective ICC with Cartel Prices below Monopoly Level (Vitamin C)](image)

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

Theoretically speaking, our sensitivity analysis in Appendices E.2 and E.3 leaves open the question of whether both reallocating quotas and changing total supply (in order for all firms’ ICCs to hold) would have been possible. Even if it had been, however, firms would not necessarily have been able to achieve it in practice. As section 2.1 described, the cartelists were wary of any renegotiation. Moreover, any small improvement in the ICC is likely to be dominated by a fixed cost of extra communication and/or the likelihood of paying penalties.62

---

61 Under static Nash, the cartel firms’ collective output surpasses its perfectly coordinated (i.e., monopoly) level by approximately 30% throughout the sample period. Hence, the ICC would hold trivially at collective output levels near “130%.”

62 According to Bos, Davies, Harrington, and Ormosi (2017), if shutting down the cartel reduces the
E.4 Monitoring Lag

We set the monitoring lag at $L = 3$ because the cartel met every three months to cross-check self-reported sales records with the official trade statistics. Did they really have to meet so frequently? What if the governments had accelerated or delayed the release of statistics? Figure 22 shows $L = 1$ (more timely information) facilitates collusion, whereas $L = 12$ (delayed release of information) makes the ICC more stringent. Thus, a monitoring lag affects cartel stability as the theory predicts, and the difference between quarterly and annual lags is quantitatively important.

![Figure 22: Collective ICC under Alternative Monitoring Lags (Vitamin C)](image)

Note: All values are multiplied by $1 - \beta$ (i.e., rescaled as the average period profits), with $\beta = 0.85$.

E.5 Rational, Static, and Adaptive Expectations

Figure 23 shows results under an alternative assumption that the cartel firms held adaptive (instead of rational) expectations regarding the demand shifter $\tilde{X}_t$. We specify that firms extrapolate future $\tilde{X}_t$ based on its 13 most recent observations (i.e., with a one-year lagging window), which makes their demand forecast highly sensitive to the latest observations (left panel). The corresponding ICC is binding in most years, because expected demand growth is too low in 1991 and 1996, and high-cost firms do not necessarily anticipate increased profits even when they foresee higher demand in 1992–1995 (right panel). Thus, adaptive expectations do not appear to match this empirical context.

We also considered a hypothetical scenario in which the cartel rationally expects the high level of fringe supply from China in 1995 from the beginning. If the firms had correctly foreseen the likelihood of paying penalties, a lower bound on the extent of collusion will exist because the benefits from collusion must be sufficient to offset the expected penalties.

The reader might wonder why the ICCs of high-cost firms would be violated under the expectation of a spectacular demand growth. The reason is that firms are asymmetric and the cartel’s quota system is based on the 1990 market shares. When $\tilde{X}_t$ doubles or triples, cost asymmetry becomes relatively less important in determining Cournot-Nash market shares (see section 5.2), which means high-cost firms will be increasingly dissatisfied with their original quotas.
eventual dominance of the Chinese firms, some of them would have foreseen practically zero profit in the late 1990s, and the ICC between 1991 and 1994 would have been as stringent as that in 1995 and 1996. These pictures are difficult to reconcile with the fact that the cartel successfully operated until 1995.

### E.6 Log-Linear Demand

Consider the following log-linear demand:

$$
\log Q^D_t = \alpha_0' + \alpha_1' \log P_t + \alpha_2' \log X_t + \varepsilon'_t,
$$

where $\alpha_1' = \frac{d \log Q}{d \log P}$ can be interpreted as the price elasticity of demand. Regardless of the functional form of the demand, the Cournot FOC remains unchanged and identifies $\frac{dP}{dQ}$. Hence, the identification strategy does not depend on this choice, and Table 10 shows marginal-cost estimates that are similar to the baseline version in Table 2.

By contrast, the effective demand shifter is recovered as the residual of the demand function,

$$
\bar{X}'_t \equiv \alpha_0' + \alpha_2' \log X_t + \varepsilon'_t
$$

and hence it depends on the functional form by construction in a more important manner. Figure 24 shows the effective demand shifter under the log-linear specification exhibits a catastrophic downturn in 1996 and never recovers. Its level in the late 1990s is as low as that in the early 1980s. This feature is at odds with the testimony of industry experts, which mentioned a secular growth trend of demand. For these reasons, we prefer our baseline, linear demand specification.
Table 10: Estimates of Log-Linear Demand and Costs (Vitamin C)

<table>
<thead>
<tr>
<th>Imputation method</th>
<th>Equal split</th>
<th>EM algorithm</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(1)</td>
<td>(2)</td>
</tr>
<tr>
<td>Price-elasticity of demand</td>
<td>-0.983 (0.001)</td>
<td>-0.982 (0.001)</td>
</tr>
<tr>
<td>$\gamma_{takeda}$</td>
<td>3.276 (0.063)</td>
<td>3.264 (0.067)</td>
</tr>
<tr>
<td>$\gamma_{e.merck}$</td>
<td>4.470 (0.050)</td>
<td>4.468 (0.051)</td>
</tr>
<tr>
<td>$\gamma_{basf}$</td>
<td>4.929 (0.086)</td>
<td>4.928 (0.087)</td>
</tr>
<tr>
<td>Number of observations</td>
<td>112</td>
<td>112</td>
</tr>
</tbody>
</table>

Note: Standard errors in parentheses are based on 1,000 block-bootstrap samples, where each block consists of 12 consecutive months of a calendar year. In column 1, we divide each annual observation of firm-level output by 12 to impute monthly outputs. In column 2, we use the EM algorithm to impute monthly outputs as part of the estimation procedure.

Figure 24: Log-Linear Demand Estimates

E.7 Long-term Contracts

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.\(^{64}\) 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
\[
\frac{1}{12} P_t q_{i,t} + \frac{1}{12} \beta P_t q_{i,t} + \cdots + \frac{1}{12} \beta^{11} P_t q_{i,t} = \frac{1 - \beta^{12}}{1 - \beta} P_t q_{i,t}.
\]
Because \( \frac{1 - \beta^{12}}{1 - \beta} \) 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
\[
1 - \beta^{12} \frac{\pi_i|t}{1 - \beta} + \sum_{k=1}^{\infty} \beta^{12k} \frac{1 - \beta^{12}}{1 - \beta} \frac{\pi_i|t+12k}{1}
\]
\[
\beta \left( 1 - \beta^{12} \frac{\pi_i|t+1}{1 - \beta} + \sum_{k=1}^{\infty} \beta^{12k+1} \frac{1 - \beta^{12}}{1 - \beta} \frac{\pi_i|t+1+12k}{1} \right)
\]
\[
\cdots + \beta^{11} \left( 1 - \beta^{12} \frac{\pi_i|t+11}{1 - \beta} + \sum_{k=1}^{\infty} \beta^{12k+11} \frac{\pi_i|t+11+12k}{1} \right),
\]
whereas firm \( i \)'s payoff under non-compliance is
\[
1 - \beta^{12} \frac{\pi_i|t}{1 - \beta} + \sum_{k=1}^{\infty} \beta^{12k} \frac{1 - \beta^{12}}{1 - \beta} \frac{\pi_i|t+12k+1}{1}
\]
\[
\beta^2 \left( 1 - \beta^{12} \frac{\pi_i|t+1}{1 - \beta} + \sum_{k=1}^{\infty} \beta^{12k+1} \frac{1 - \beta^{12}}{1 - \beta} \frac{\pi_i|t+2+12k}{1} \right)
\]
\[
\beta^3 \left( 1 - \beta^{12} \frac{\pi_i|t+2}{1 - \beta} + \sum_{k=1}^{\infty} \beta^{12k+2} \frac{1 - \beta^{12}}{1 - \beta} \frac{\pi_i|t+3+12k}{1} \right)
\]
\[
\cdots + \beta^{11} \left( 1 - \beta^{12} \frac{\pi_i|t+11}{1 - \beta} + \sum_{k=1}^{\infty} \beta^{12k+11} \frac{\pi_i|t+11+12k}{1} \right).
\]
Rearranging, the equilibrium condition is
\[
\sum_{k=1}^{\infty} \beta^{k} \pi_i|t+k-1 \geq \sum_{k=1}^{3} \beta^{k} \pi_i|t+k-1 \geq \sum_{k=1}^{\infty} \beta^{k} \pi_i|t+k-1
\]
for each \( i \) and \( \tau \geq t \). That is, \( \min_{i,\tau \geq t} (V^C_{i,\tau|t} - V^D_{i,\tau|t}) \geq 0 \), as in (13).

\(^{64}\)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 period \( t \) through \( t + 11 \), \( \frac{q_{i,t}}{12} \) for each month. Then \( q_{i,1} = q \) is delivered over period 1 through 12, \( q_{i,2} = q \) is delivered over period 2 through 13, and so on. Hence, the total quantity delivered is \( q \) per month in the steady state.
Appendix F: Merger Simulation

F.1 Formal Analysis of Mergers’ Impact on the ICC

Suppose $n$ firms face the inverse demand,

$$P = a - b \sum_j q_j,$$

(34)

and compete in quantities. Firm $i$’s marginal cost is

$$c_i = c_i^* + \gamma_i,$$

(35)

where $i^*$ is the low-cost firm and the cartel leader (e.g., Roche), and $\gamma_i$ is the constant gap between their costs.

**Cournot with Asymmetric Costs** Firm $i$’s FOC in Cournot competition is

$$\frac{d}{dq_i} \left( a - b \sum_j q_j - c_i \right) q_i = 0,$$

(36)

or

$$q_i = \frac{a}{b} - \sum_j q_j - \frac{c_i}{b}.$$  

(37)

Adding up (37) across $n$ firms and rearranging, we have

$$\sum_i q_i = \frac{n}{n+1} \frac{a}{b} - \frac{1}{n+1} \sum_i c_i b = \frac{n}{n+1} \frac{a}{b} - \frac{n \bar{c}}{n+1 b'},$$

(38)

where $\bar{c} = \frac{1}{n} \sum_i c_i$ is the average of all firms’ marginal costs. Substituting (38) into (36) yields

$$q_i = \frac{1}{n+1} \frac{a}{b} + \frac{n}{n+1} \bar{c} - \frac{c_i}{b}.$$  

(39)

Hence, firm $i$’s market share in static Nash equilibrium is

$$s_i = \frac{q_i}{\sum_j q_j} = \frac{\frac{1}{n+1} \frac{a}{b} + \frac{n}{n+1} \bar{c} - \frac{c_i}{b}}{\sum_j q_j},$$

(40)

the price is

$$P = a - b \left( \frac{n}{n+1} \frac{a}{b} - \frac{n}{n+1} \bar{c} \right) = \frac{1}{n+1} a + \frac{n}{n+1} \bar{c},$$

(41)
and the static Nash profit is
\[ \pi_i^N = \left( \frac{1}{n+1} a + \frac{n}{n+1} \bar{c} - c_i \right) \left( \frac{1}{n+1} \frac{a}{b} + \frac{n}{n+1} \bar{c} - \frac{c_i}{b} \right). \] (42)

**Collusion with Asymmetric Costs**

From section 4.4, the cartel’s collective output is
\[ Q^C = \frac{a - c_i^*}{2b}, \] (43)
and firm \( i \)'s collusive profit is
\[ \pi_i^C = s_i \frac{a - c_i^*}{2b} \left( \frac{a - c_i^*}{2b} - \gamma_i \right), \] (44)
where its output quota is based on \( s_i \) in (40).

**Deviation Gain with Asymmetric Costs**

Suppose firm \( i \) maximizes the deviation gain,
\[ \pi_i^D = \max_{q_i} \left( a - bq_i - (1 - s_i) \frac{a - c_i^*}{2} - c_i - \gamma_i \right) q_i. \] (45)
Its optimal deviation quantity is
\[ q_i^D = \frac{(a - c_i^*) (1 + s_i) - 2 \gamma_i}{4b}. \] (46)

**Incentive Compatibility**

Firm \( i \)'s deviation gain is
\[ (1 + \beta + \beta^2) (\pi_i^D - \pi_i^C) = (1 + \beta + \beta^2) \left( a - b \frac{(a - c_i^*) (1 + s_i) - 2 \gamma_i - (1 - s_i) \frac{a - c_i^*}{2} - c_i - \gamma_i}{4b} \right) \]
\[ \times \left( a - c_i^* \right) \left( 1 + s_i \right) - 2 \gamma_i - s_i \frac{a - c_i^*}{2b} \left( \frac{a - c_i^*}{2b} - \gamma_i \right), \] (47)
and the deviation loss is
\[ \frac{\beta^3}{1 - \beta} (\pi_i^C - \pi_i^N) = \beta^3 \left( \frac{s_i}{n+1} a + \frac{n}{n+1} \bar{c} - c_i \right) \left( \frac{1}{n+1} \frac{a}{b} + \frac{n}{n+1} \bar{c} - \frac{c_i}{b} \right), \] (48)
where
\[ s_i = \frac{1}{n+1} a + \frac{n}{n+1} \frac{\bar{c} + \gamma_i}{b} - \frac{c_i + \gamma_i}{b}. \] (49)

A merger reduces \( n \) and could alter \( \gamma_i \) of the merging firm. Various combinations of pre- and post-merger configurations, \((n, \gamma_i)\) and \((n-1, \gamma_i')\), exist such that the cutoff discount factor is larger
under \((n - 1, \gamma_i')\) despite the smaller number of firms. Hence, collusion could become more difficult when a merger affects both \(n\) and \(\gamma_i\).

### F.2 Counterfactual ICCs

Sections 7.3 and 7.4 show summaries of various merger counterfactuals in terms of the cumulative survival frequencies of the cartel. Figure 25 shows the ICC that underlie the BASF-Takeda merger counterfactuals (Figures 10 and 11). A comparison of the top panels suggests this merger (without synergy) could have relaxed the ICC, but the improvement does not seem sufficient to save the cartel from its collapse in 1995 or 1996. The middle panels show further improvements in these critical years and suggest mild synergy \((\sigma = 0.05–0.10)\) might help. However, the ICC in earlier years becomes more stringent at the same time. The bottom panels highlight this problem and show larger synergy \((\sigma = 0.05–0.10)\) could increase asymmetry and destabilize the cartel. Thus, these time profiles of the ICC suggest efficiency gains may lead to negative coordinated effects at various points in time.

Figure 26 shows the ICC under the six different merger simulations in Table 3 and Figure 12. Whereas the previous figure exhibited some tradeoff in cartel stability between the early 1990s and the late 1990s, the six plots in this figure break such patterns and feature larger shifts in the overall level of ICC. The only mergers that unambiguously facilitate collusion throughout the sample period are Mergers #1 and #4. These two mergers make the cost profile more symmetric by eliminating the least competitive firm(s) such as BASF. By contrast, the four other mergers make the ICC significantly more stringent by eliminating firms “in the middle” in terms of cost competitiveness (i.e., Takeda and/or E. Merck), which leads to more polarized cost profiles. Thus, despite making market structure more concentrated, many mergers turn out to have negative coordinated effects.
Figure 25: Counterfactual ICCs under BASF-Takeda Merger by Synergy Level

Note: All values are multiplied by $1 - \beta$ (i.e., rescaled as the average period profits), with $\beta = 0.85$. 

A-34
Figure 26: Counterfactual ICCs under Six Different Mergers

Note: All values are multiplied by $1 - \beta$ (i.e., rescaled as the average period profits), with $\beta = 0.85$. 
References


A-40


