The Ice Cream Split: Empirically Distinguishing Price and Product Space Collusion

Christopher J. Sullivan*
University of Michigan

Date Revised: July 8, 2017
Click Here for the Latest Version

Abstract

In a market for differentiated products, firms have the ability to collude on the choice of products offered in addition to or in lieu of colluding on the prices charged for those products. However, the empirical literature has only considered price collusion. This paper proposes a methodology to measure product space and price collusion. To do so, I model firms as competing in an infinitely repeated extensive form game. I show that a subset of equilibria to the dynamic game can be represented as subgame perfect Nash equilibria to a static game. The static equilibria index the degree to which firms collude through reduced-form parameters. I show that these parameters can be estimated using standard techniques when researchers have access to market level data. I then use this methodology to study competition in the market for super-premium ice cream during 2013. I find evidence that Ben & Jerry's and Häagen-Dazs not only colluded on the prices they charged, but substantially colluded in the choice of flavors that were offered. Then, I construct counterfactuals to measure the impact of collusion on the set of products offered, the prices charged, and welfare. I find that conventional policy analysis that considers only price collusion understates the welfare effects.

*Department of Economics, University of Michigan, 611 Tappan Street, Ann Arbor, MI 48109; sullivcj@umich.edu
This paper is a revised version of the first and third chapters of my 2017 doctoral dissertation at the University of Michigan. I would first like to thank my advisor, Daniel Ackerberg for the incredible amount of time, energy, and support he has given to me. The other members of my committee: Ying Fan, David Miller, and Yesim Orhun provided tremendous advice and encouragement for which I can't thank them enough. This paper was vastly improved through conversations with Jim Adams, Sarah Johnston, Paul Brehm, Adam Dearing, Alan Griffith, Julian Hsu, Andrew Litten, Andrew Usher, and by comments from seminar participants at the University of Michigan. I would also like to thank my grandmother who always kept a pint of Chunky Monkey in her freezer for me. All remaining errors are my own.

"When the smooth get chunky, the chunky get smooth." - Ben Cohen, co-founder of Ben & Jerry's

1 Introduction

At the heart of oligopoly lies the tension between adhering to or defecting from a collusive agreement, be it tacit or explicit. Because collusion is illegal and actively prosecuted, there is an exhaustive empirical literature trying to detect collusion and measure its impact on welfare. However, these papers only consider collusion on price. In a market for differentiated products, might not firms collude on the types of products offered in addition to or in lieu of colluding on the prices charged for those products? As was noted in Fershtman and Pakes (2000) "... the ability to collude will have an impact on the variety, cost, and quality of the products marketed by the industry, and this can have as much or more of an effect on welfare as do the price effects of collusion" (pg. 208).

Understanding product space collusion is important for many reasons, two of which I highlight here. First, economists and antitrust authorities may be ignoring many instances of anticompetitive behavior by restricting their attention to price. Consider a market with two fixed products. In one state of the world, each firm offers both products and they agree to collude on price. In another state of the world, the firms compete on price but do so after agreeing to produce different products. With our current techniques, collusion would only be detected in the first scenario, yet both are anticompetitive, leading to higher prices for consumers and lower social welfare.

Second, if we relax the assumption that the set of products remain fixed, then collusion may affect the positioning of a firm's offerings in the product space. When one allows for this possibility, the effect of increased market power on consumer surplus becomes ambiguous, as has been discussed in the literature studying endogenous product selection and mergers.[1] In particular, if product and price collusion result in firms offering a more diverse set of products, this increased variety may lead to higher consumer welfare and offset the loss in consumer surplus from higher prices.

In this paper, I propose an estimation strategy that allows researchers to separately test for both product and price collusion. Unlike previous empirical studies of collusion, I explicitly derive the reduced-form parameters measuring collusion from the structural parameters of a repeated game. I show that these reduced-form parameters can be estimated in the setting familiar to IO economists. I then apply this methodology to study competition in the market for super-premium ice cream in 2013. In doing so, I find evidence that Ben & Jerry's and Häagen-Dazs colluded both in the set of flavors they offered and the prices they charged for those flavors. Then, I construct counterfactuals to measure the impact of collusion on the set of products offered, the prices charged, and welfare. With these counterfactuals, I am able to measure the degree to which ignoring collusion on the product space misstates the impact of collusion on outcomes and welfare. I find that ignoring product collusion understates the negative effect on welfare.

If the antitrust authorities do not fully account for collusion by ignoring the product space, current antitrust policies could have unintended effects. Consider a scenario in which firms are

[1] These include Draganska, Mazzeo and Seim (2009) and Fan (2013).

colluding on both the choice of products and price. If the antitrust authorities detect price collusion in a market and take steps to prevent it, firms also colluding in the product space might respond by adjusting their product offerings. This repositioning in the product space can lead to unintended welfare consequences; in particular, the policy intervention may result in lower consumer surplus. To investigate this possibility, I simulate a counterfactual in which Ben & Jerry's and Häagen-Dazs continue to collude on product choice but are prevented from colluding on price. Though the counterfactual is not robust to the Lucas Critique and is therefore speculative, it is offered as an illustration of the potential ramifications that might result from implementing policies without considering the full scope of firm collusion. In doing so, I find that policy interventions prohibiting only price collusion have the potential to reduce consumer surplus relative to policy inaction. This occurs because the firms choose to greatly curtail their product offerings when they are limited to only colluding in the product space.

The market for super-premium ice cream provides a setting conducive to investigating the role of product space collusion. The market has been dominated by two firms: Ben & Jerry's and Häagen-Dazs. These brands have long been associated with distinct styles of ice cream: Häagen-Dazs is known to produce "smooth", traditional flavors[2] while Ben & Jerry's sells so-called "chunky" flavors: ice cream to which extra ingredients like chocolate, caramel, candy, and baked goods have been added.[3]

There is anecdotal evidence suggesting that Ben & Jerry's and Häagen-Dazs have coordinated their product choices. First, it seems unlikely that the assortments offered by the brands qualify as best responses. Consider the quintessential flavors: chocolate, coffee, strawberry, and butter pecan. These popular flavors were responsible for 28.2 percent of Häagen-Dazs sales in 2013, so it should come as no surprise that Ben & Jerry's also makes them. However, Ben & Jerry's only sells them at its ice cream shops or as the bases for chunky flavors. Secondly, the brands appear to have engaged in periodic product space wars, consistent with collusive behavior supported by trigger strategies as predicted by Green and Porter (1984) and Rotemberg and Saloner (1986). *The New York Times* covered the first product space war between the brands in 1994.

> *"Yet, by all accounts, the slower growth in the super-premium market has meant tougher competition...[Häagen-Dazs and Ben & Jerry's] have invaded each other's turf...The fight began two years ago [1992] when Häagen-Dazs, long known for its 'smooth' ice cream, went after Ben & Jerry's market by introducing its own versions [of] 'chunky' ice cream, dubbed 'Exträas'...Last spring Ben & Jerry's...retaliated with its own smooth varieties. To promote its smooth flavors, the company is advertising on television for the first time...Häagen-Dazs isn't amused."*

[2]The top 6 best selling flavors for Häagen-Dazs in 2013 were vanilla, chocolate, coffee, strawberry, butter pecan, and vanilla bean.

[3]Ben & Jerry's top 6 best selling flavors in 2013 were Cherry Garcia (cherry ice cream with fudge chunks and cherry pieces), Half Baked (blend of chocolate chip cookie dough and chocolate fudge brownie), Chocolate Fudge Brownie, Chocolate Chip Cookie Dough, Coffee Heath Bar Crunch, and Chunky Monkey (banana ice cream with walnuts and dark chocolate chunks).

Describing his company's strategy to the media during that period, Ben Cohen asserted "When the smooth get chunky, the chunky get smooth."

In order to credibly claim that the observed pattern of product assortment is indicative of collusive behavior, I have to rule out confounding explanations, chief amongst them that consumer preferences are responsible for firm differentiation[4] and that the firms have cost differences that manifest themselves in different product choices. I am able to explicitly control for these stories. Specifically, I estimate demand using the Berry, Levinsohn and Pakes (1995) algorithm while including a fixed effect for each brand-flavor combination as in Nevo (2001). Thus, I control for average quality and perception differences between the versions of flavors produced by each brand. Secondly, I model firms as sequentially choosing products and prices. This permits me to estimate marginal cost using the first order conditions of each firm's second stage pricing decision, allowing me to control for cost differences when estimating the product collusion parameters in the first stage. As such, estimated values of the reduced-form parameters indicative of collusion should be viewed as robust to these two alternative explanations.

The empirical literature largely models collusion using reduced-form parameters in a static game in lieu of specifying a structural model of collusion. Using simpler static models to proxy for the dynamic structural game allows researchers to avoid the computational issues associated with estimating multi-agent dynamic games with multiple equilibria. One strand of the empirical literature measures collusion by estimating reduced-form profit weights. These papers, which include Bresnahan (1987), Nevo (1998), and Miller and Weinberg (forthcoming), model firms as maximizing a weighted sum of their own profit and the profit of each of their rivals. The weight a firm places on each of its rivals captures the degree to which it internalizes the impact of its actions on that rival. As will be seen in Section 2, I extend the reduced-form framework used in these papers to a two-stage static game in which firms sequentially choose products and prices.

A criticism of the literature estimating reduced-form parameters that measure firm conduct is that these parameters are not derived from a structural model, making interpretation of the collusion parameters difficult. Furthermore, it is unclear how the reduced-form parameters might change in alternative market or policy settings without mapping the structural parameters to the reduced-form parameters. In an unpublished paper, Black, Crawford, Lu, and White (2004) (hereafter BCLW) show how the structural parameters in a repeated Bertrand game can be mapped to the reduced-form profit weights measuring price collusion. In Section 2, I extend their framework to a model in which firms compete in two stages, a product setting stage and a price stage.

My research is also related to Hackner (1995) and Xu and Coatney (2015), which to my knowledge are the only two theoretical papers that explicitly model collusion on the product space. Furthermore, because, the counterfactuals allow me to measure the effects of collusion on the set of products offered, this paper relates to the literature on endogenous product choice. In addition to Draganska, Mazzeo and Seim (2009) (hereafter DMS), which considers the market for premium ice

[4]There could be several reasons for this. First, each brand could specialize in flavors which consumers think are higher quality than the version its rival could produce. Secondly, though no quality difference in each brand's version of a flavor exists, consumers may have come to strongly associate each brand with a style of ice cream.

cream, and Fan (2013), these include Mazzeo (2002), Crawford and Shum (2007), Sweeting (2013), Crawford and Yurukoglu (2012), Orhun (2013), Orhun, Venkataraman and Chintagunta (2015), Fan and Yang (2016), and Wollmann (2016).

The paper proceeds as follows: Section 2 presents the repeated game and derives its reduced-form representation. Section 3 provides background on the market for super-premium ice cream. Section 4 presents the empirical model to be estimated. Section 5 describes the data. Section 6 discusses identification and estimation of the parameters in the empirical model. Results and counterfactuals are presented in Section 7 and 8 respectively. Section 9 discusses extensions and concludes.

2 Structural Model

As was mentioned above, estimating the structural parameters in a dynamic model of collusion is computationally difficult. Therefore, I begin in Section 2.1 by presenting a simpler reduced-form game, which will form the basis of the empirical model. The reduced-form game is a static two-stage game that contains two parameters for each firm, one indexing the degree to which that firm colludes in the choice of products and one indexing the degree to which it colludes in price. However, it is difficult to interpret the reduced-form parameters in the absence of a structural model. I address this concern in Section 2.2, where I show the subgame perfect Nash equilibrium in the reduced-form game is equivalent to a stationary equilibrium of an infinitely repeated game. To show this equivalence, I extend the methodology presented in BCLW (2004) to two stages. Because the equilibria are equivalent, I can derive a mapping from the structural parameters of the repeated game to the reduced-form parameters. For simplicity of exposition, the first two sections only consider one market. In Section 2.3, I extend the results to a setting in which firms compete in multiple markets. As will be seen, competition in multiple markets allows me to consider reduced-form collusion parameters that are robust to the Corts Critique.

2.1 The Reduced-Form Game

I consider a reduced-form game that is related to the strain of the IO literature, e.g. Bresnahan (1987), Nevo (1998), and Miller and Weinberg (forthcoming), that models price collusion using the following approach. Two firms,[5] $b = \{1, 2\}$ compete in a static Bertrand game, each choosing price p_b to maximize a weighted sum of its own profit and the profit of its rival:

$$\max_{p_b} \pi_b + \theta_b \pi_{-b} \tag{1}$$

The Nash equilibrium is a set of prices satisfying the first order condition for each firm.

$$\frac{\partial \pi_b}{\partial p_b} + \theta_b \frac{\partial \pi_{-b}}{\partial p_b} = 0 \tag{2}$$

[5] These papers consider N firms. I consider only two firms since, in my empirical example, there are only two firms: Ben & Jerry's and Häagen-Dazs.

The reduced-form parameters $\theta = (\theta_b, \theta_{-b})$ are then estimated via the first order conditions. There are several advantages to this approach. Foremost, the first order conditions index a large set of equilibria, from Nash-Bertrand when $\theta = 0$ to perfect collusion when $\theta = 1$. Furthermore, the θ's have a compelling interpretation: they summarize the degree to which each firm internalizes the externality its pricing decision imposes on its rival, which is consistent with their role in measuring collusive behavior. In addition, the θ's are estimated via standard techniques.

Given the appeal of this framework, I use it as a basis for modelling product and price collusion. Specifically, I extend this framework to a two-stage game in which firms first choose products, then choose prices. Formally, each product represents a unique combination of K continuous product characteristics such that the product space is given by $\mathbb{R}^K$. I let $d_b \subset \mathbb{R}^K$ be the set of products offered by firm b where $d_{bj} = (d_{bj1}, \ldots, d_{bjK})$ represents the K-dimensional vector of characteristics defining the j^{th} product in d_b. In the second stage, after observing the product offerings of their rival, firms simultaneously set the prices for these products. The price charged by firm b for product j is $p_{bj}(d_b, d_{-b})$ which depends on the entire set of products offered by both firms. At the end of the two stages, the firm receives profits according to

$$\pi_b = \tilde{\pi}_b(d_b, d_{-b}, p_b, p_{-b}) - \eta_b N_b \tag{3}$$

where $\tilde{\pi}_b$ are the variable profits earned by the firm, η_b is a per product fixed cost, and N_b is the number of products offered by the firm.[6] I assume that the variable profits for each firm are concave in the prices and characteristics chosen by that firm.

For the two stage game, the firms maximize a weighted sum of their own profits in each stage where firm b places a weight $\theta_{1,b}$ on its rival's profit in the product stage and weight $\theta_{2,b}$ in the pricing stage. $\theta_{1,b}$ and $\theta_{2,b}$ are allowed to differ across the stages. The subgame perfect Nash equilibrium (SPNE) to this game can be found by backwards induction. In the second stage, firm b chooses its price given the set of products being offered to solve:

$$\max_{p_b} \ \pi_b\Big(d_b, d_{-b}, p_b, p_{-b}\Big) + \theta_{2,b}\pi_{-b}\Big(d_b, d_{-b}, p_b, p_{-b}\Big) \tag{4}$$

Thus, for a given choice of products, the equilibrium set of prices satisfies the first order condition for each firm.

$$\frac{\partial \pi_b}{\partial p_b} + \theta_{2,b}\frac{\partial \pi_{-b}}{\partial p_b} = 0 \qquad\qquad \forall b \tag{5}$$

In the first stage, firm b chooses a subset of products to offer taking into account the second stage pricing decision.

$$\max_{d_b} \ \pi_b\Big(d_b, d_{-b}, p_b(d_b, d_{-b}), p_{-b}(d_b, d_{-b})\Big) + \theta_{1,b}\pi_{-b}\Big(d_b, d_{-b}, p_b(d_b, d_{-b}), p_{-b}(d_b, d_{-b})\Big) \tag{6}$$

[6] The restricted specification where fixed costs do not depend on product characteristics is motivated by my empirical context, super-premium ice cream. This is discussed further in Section 4.2.

Since the characteristics defining each product are continuous, a subset of the necessary conditions for the equilibrium set of products can generally be written as:[7]

$$\frac{\mathbf{d}\pi_b}{\mathbf{d}d_b} + \theta_{1,b}\frac{\mathbf{d}\pi_{-b}}{\mathbf{d}d_b} = 0 \qquad\qquad \forall b \tag{7}$$

where the total derivatives in the first order condition are defined as:

$$\frac{\mathbf{d}\pi_b}{\mathbf{d}d_b} = \frac{\partial\pi_b}{\partial d_b} + \frac{\partial\pi_b}{\partial p_b}\frac{\partial p_b}{\partial d_b} + \frac{\partial\pi_b}{\partial p_{-b}}\frac{\partial p_{-b}}{\partial d_b} \tag{8}$$

and

$$\frac{\mathbf{d}\pi_{-b}}{\mathbf{d}d_b} = \frac{\partial\pi_{-b}}{\partial d_b} + \frac{\partial\pi_{-b}}{\partial p_b}\frac{\partial p_b}{\partial d_b} + \frac{\partial\pi_{-b}}{\partial p_{-b}}\frac{\partial p_{-b}}{\partial d_b} \tag{9}$$

In this reduced-form game there is a distinct price collusion parameter and product collusion parameter for each firm. At first it may seem odd to the reader that a firm has a different collusion parameter in each stage of the game, changing the objective function in each stage. I offer two related justifications. Intuitively, it seems likely that a firm would differentially internalize the effect of its product and pricing decisions on its rival. In the next section, I also provide a theoretical justification by deriving a mapping from the structural parameters in the underlying repeated game to the reduced-form parameters. In order to derive the mapping, I extend the work of BCLW (2004). In BCLW (2004), the authors show how to derive the mapping for a repeated Bertrand game in which firms collude in price. I adapt their framework to accommodate a two-stage game in which firms can compete in prices and products.

2.2 The Repeated Game

In this section, I present a structural model of collusion in a market for differentiated products. The model is similar in spirit to that of Rotemberg and Saloner (1986). Two firms with perfect information and a common discount factor play an infinitely repeated extensive-form game. Every period has two stages. During the first stage, each firm simultaneously chooses a set of products to produce. In the second stage, after observing the product choices of their rival, firms simultaneously set prices. Firms receive one payoff at the end of the period. At the start of each period, before choosing products or prices, both firms observe the realization of a vector of excluded, exogenous, random demand shifters, x_t. These realizations are allowed to be arbitrarily correlated over time, thus generalizing the model in BCLW (2004) which can be interpreted as implicitly including demand shocks that are perfectly correlated over time. Explicitly including exogenous demand shocks in the model is useful as much of the literature follows Bresnahan (1982) in using exogenous demand

[7] These conditions would not necessarily hold if a firm was indifferent between the number of products to offer. However, because the characteristics are continuous, indifference is unlikely. Therefore, these conditions generically hold at the equilibrium product choice. In addition, I omit the first order condition associated with the optimal number of products because that is a discrete choice. These caveats hold throughout this section.

shifters and rotators to identify reduced-form conduct parameters. Likewise, I rely on instruments that shift and rotate the demand curve to identify the reduced form price collusion parameters in Section 6.2.

The strategy space and profits in each period are identical to those defined in the reduced-form game. In period t, firm b chooses a set of products $d_{bt} \in \mathbb{R}^K$ to offer and sets a price $p_{bt}(d_{bt}, d_{-bt})$. The profits earned by firm b at the end of the period are given as:

$$\pi_{bt}(d_{bt}, d_{-bt}, p_{bt}, p_{-bt}; x_t) = \tilde{\pi}_{bt}(d_{bt}, d_{-bt}, p_{bt}, p_{-bt}; x_t) - \eta_b N_{bt} \tag{10}$$

where I assume that the variable profits for each firm are concave in the prices and characteristics chosen by that firm. Each firm b maximizes the present discount value of its future stream of profits $\sum_{t=0}^{\infty} \delta^t \pi_{bt}$ where δ is a discount factor common to both firms.

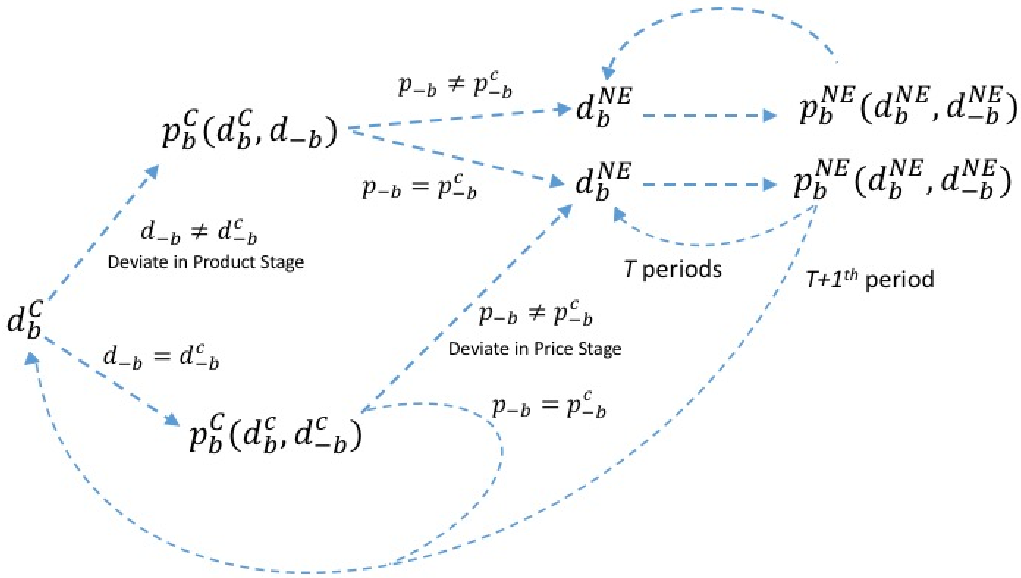
The Folk Theorem makes it clear that this game has a large set of equilibria. I focus attention on on finding and describing one particular stationary equilibrium which is equivalent to the subgame perfect Nash equilibrium of the reduced-form game presented in the previous section. I impose two refinements in order to select the desired equilibrium. First I define a set of stationary strategies which ensure that the selected equilibrium will be a subgame perfect Nash equilibrium for the firms in the repeated game. These strategies combine both stick and carrot and grim trigger punishments to keep the firms on the equilibrium path. However, these strategies still permit a large set of equilibria. Thus, I make an additional refinement: firm actions must be Pareto optimal in that no one firm can be made better off without hurting its rival. Taken together, these refinements select the appropriate equilibrium.

2.2.1 Firm Strategies

Stationarity requires that, in every period t, firm b chooses products d_{bt} and prices p_{bt} according to product and pricing functions defined over the demand shocks x. Thus, for all periods t, $d_{bt} = d_b(x_t)$ and $p_{bt} = p_b(d_1(x_t), d_2(x_t), x_t)$. Before play begins, colluding firms agree to a specific product function such that in period t, the collusive products are given by $(d_{1t}^C, d_{2t}^C) = (d_1^C(x_t), d_2^C(x_t))$. Likewise, firms agree to a specific collusive pricing function such that in period t the pair of collusive prices is defined as: $\Bigg(p_1^C(d_{1t}, d_{2t}; x_t), p_2^C(d_{1t}, d_{2t}; x_t)\Bigg)$. Because the firms could collude on just the pricing stage, there is a prescribed collusive price for every possible choice of products, not just for the collusive products. Then play proceeds according to the strategies depicted in Figure 1.[8] Firm b begins the game offering d_{bt}^C in the product stage. If firm b observes its rival produced d_{-bt}^C, it charges $p_b^C(d_{1t}^C, d_{2t}^C; x_t)$ in the pricing stage. If its rival does not defect in the pricing stage, firm b begins the next period by again offering the collusive products. The two firms continue cooperating until defection is observed.

[8]These strategies are non-standard; admittedly both simpler and harsher punishments exist. However, if firms are patient enough, these strategies are sufficient to maintain any level of collusion. Also, it would be difficult to derive the optimal penal code for this extensive form game. See Mailath, Nocke and White (forthcoming) which explains the difficulties in doing so.

Figure 1
Diagram Of Equilibrium Strategies For Firm b



There are three ways in which a firm can defect in a given period: in the product stage only, in the pricing stage only, or in both the product stage and pricing stage. If any of these cases occur, the punishment phase begins at the start of the next period. During the punishment phase, firms revert to playing the actions associated with the one-shot subgame perfect Nash equilibrium $d_{bt}^{NE} = d_b^{NE}(x_t)$ and $p_{bt}^{NE} = p_b^{NE}(d_{bt}^{NE}, d_{-bt}^{NE}; x_t)$. These actions can be found by backwards induction. In the second stage of a period in the punishment phase, the equilibrium prices for a given set of products satisfy:

$$\frac{\partial \pi_{bt}}{\partial p_{bt}} = 0 \qquad \forall\, b \tag{11}$$

Likewise a necessary condition for firm b's choice of products is given as:

$$\frac{\partial \pi_{bt}}{\partial d_{bt}} + \frac{\partial \pi_{bt}}{\partial p_{bt}^{NE}} \frac{\partial p_{bt}^{NE}}{\partial d_{bt}} + \frac{\partial \pi_{bt}}{\partial p_{-bt}^{NE}} \frac{\partial p_{-bt}^{NE}}{\partial d_{bt}} = 0 \qquad\qquad \forall b \tag{12}$$

The length of the punishment phase depends on whether firm $-b$ deviates in only one of the stages or if it deviates in both the product and price stage. If deviation occurs in only one of the two stages, the punishment lasts for T periods after which the firms resume cooperation. However, if a firm deviates in both the product stage and price stage, the punishment phase lasts forever. Because punishment does not begin until the start of the period after which it occurs, there remains the question of what prices firms charge if deviation occurs on the product space. If firm $-b$ deviates in products, firm b offers the collusive prices given the set of products that were offered in the previous stage, represented as $p_b^C(d_{bt}^C, d_{-bt}; x_t)$. If its rival also plays the collusive prices conditional on the set of products, the punishment lasts for T periods. If firm $-b$ deviates again, the punishment lasts forever.

In order for a set of collusive actions to constitute a subgame perfect Nash equilibrium given these strategies, the present discounted value of colluding in both stages by playing these actions must exceed the present discounted value of any one-period deviation. Since there are three ways in which a firm can deviate, these strategies place three restrictions on the set of equilibria.

First, the profits from colluding in both stages must exceed the profits from deviating only in the pricing stage. This implies:

$$\pi_{bt}^{C,C} + \sum_{\tau=1}^{T} \delta^{\tau} \mathbb{E}_t \pi_{bt+\tau}^{C,C} \geq \pi_{bt}^{C,D} + \sum_{\tau=1}^{T} \delta^{\tau} \mathbb{E}_t \pi_{bt+\tau}^{NE} \tag{13}$$

where $\pi_{bt}^{C,C}$ represents the per period payoff earned by firm b from colluding in both stages, $\pi_{bt}^{C,D}$ represents the maximal profits that can be earned by firm b from colluding on the choice of products but defecting in the pricing stage, and π_{bt}^{NE} represents the one-shot subgame perfect Nash equilibrium profits earned by firm b during each period of the punishment phase. Likewise, the profits from colluding in both stages must exceed the profits from deviating in the product stage but colluding on price.

$$\pi_{bt}^{C,C} + \sum_{\tau=1}^{T} \delta^{\tau} \mathbb{E}_t \pi_{bt+\tau}^{C,C} \geq \pi_{bt}^{D,C} + \sum_{\tau=1}^{T} \delta^{\tau} \mathbb{E}_t \pi_{bt+\tau}^{NE} \tag{14}$$

where $\pi_{bt}^{D,C}$ represents the maximal profits firm b can earn by deviating in the product stage but colluding on price. Finally, the firms must find it unprofitable to deviate in both stages. This requires:

$$\pi_{bt}^{D,C} + \sum_{\tau=1}^{T} \delta^{\tau} \mathbb{E}_t \pi_{bt+\tau}^{NE} + \sum_{\tau=T+1}^{\infty} \delta^{\tau} \mathbb{E}_t \pi_{bt+\tau}^{C,C} \geq \pi_{bt}^{D,D} + \sum_{\tau=1}^{\infty} \delta^{\tau} \mathbb{E}_t \pi_{bt+\tau}^{NE} \tag{15}$$

where $\pi_{b}^{D,D}$ represents the maximal profits firm b can earn by deviating in both stages.

I assume that firms are sufficiently patient such that deviating in both stages is never profitable. Therefore, given the incentive constraints defined by (13) and (14) I can recursively find the set of actions in each stage that constitute subgame perfect Nash equilibria in period t. For any given set of products produced in the first stage, we can find the set of prices such that neither firm wants to deviate in the second stage.

$$\mathcal{P}_t(d_{1t}, d_{2t}; x_t, \delta, T) = \Bigg\{(p_{1t}, p_{2t}) \Bigg| \pi_{bt}(d_{1t}, d_{2t}, p_{1t}, p_{2t}; x_t) \\ + \sum_{\tau=1}^{T} \delta^\tau \mathbb{E}_t \pi_{bt+\tau}(d_{1t+\tau}, d_{2t+\tau}, p_1^C(d_{1t+\tau}, d_{2t+\tau}; x_{t+\tau}), p_2^C(d_{1t+\tau}, d_{2t+\tau}; x_{t+\tau}); x_{t+\tau}) \\ \geq \pi_{bt}^{C,D}(d_{1t}, d_{2t}, p_{-bt}; x_t) + \sum_{\tau=1}^{T} \delta^\tau \mathbb{E}_t \pi_{bt+\tau}^{NE}(x_{t+\tau}) \qquad \forall\, b = 1, 2\Bigg\} \tag{16}$$

Then we can find the set of products in the first stage from which firms will not want to deviate given that firms charge prices according to $\mathcal{P}_t$ in the second stage:

$$\mathcal{D}_t(x_t, \delta, T) = \Bigg\{(d_{1t}, d_{2t}) \Bigg| \pi_{bt}(d_{1t}, d_{2t}, p_{1t}^C(d_{1t}, d_{2t}; x_t), p_{2t}^C(d_{1t}, d_{2t}; x_t); x_t) \\ + \sum_{\tau=1}^{T} \delta^\tau \mathbb{E}_t \pi_{bt+\tau}(d_{1t+\tau}^C, d_{2t+\tau}^C, p_1^C(d_{1t+\tau}^C, d_{2t+\tau}^C; x_{t+\tau}), p_2^C(d_{1t+\tau}^C, d_{2t+\tau}^C; x_{t+\tau}); x_{t+\tau}) \\ \geq \pi_{bt}^{D,C}(d_{-bt}; x_t) + \sum_{\tau=1}^{T} \delta^\tau \mathbb{E}_t \pi_{bt+\tau}^{NE}(x_{t+\tau}) \qquad \forall\, b = 1, 2\Bigg\} \tag{17}$$

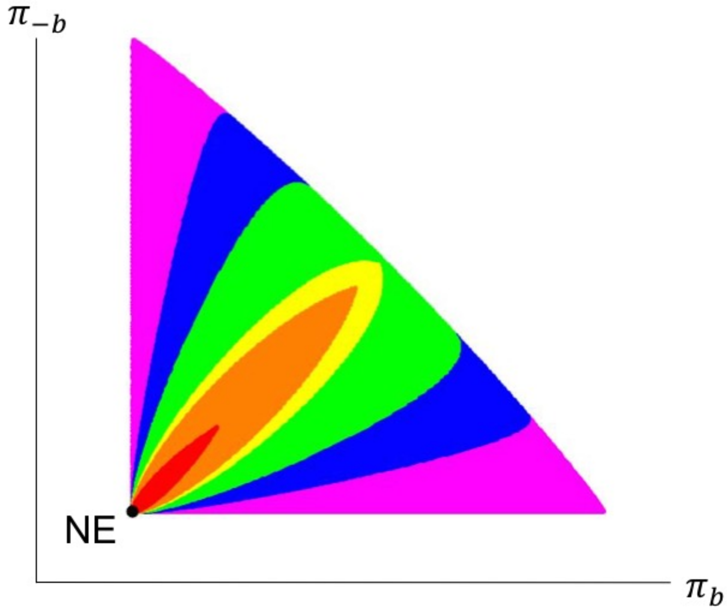
Thus a subgame perfect Nash equilibrium consists of a set of products $(d_{bt}^C, d_{-bt}^C) \in \mathcal{D}_t(x_t, \delta, T)$ and a set of prices $(p_{bt}^C, p_{-bt}^C) \in \mathcal{P}_t(d_{bt}^C, d_{-bt}^C; x_t, \delta, T)$.

Notice that for a given δ, the set of collusive payoffs is not unique. In Figure 2, I fix a choice of products (d_{bt}, d_{-bt}) and find the set of payoffs for both firms that are attainable given $\mathcal{P}_t$ and a realization of x_t.[9] When $\delta = 0$, the only feasible payoffs are those in which the firms charge the Nash prices $p_{bt}^{NE}(d_{bt}, d_{-bt})$. As δ increases above zero, so too does the set of feasible prices and payoffs. For high enough values of δ, the firms are able to achieve some payoffs on the profit possibility frontier. Because the set of equilibria is not unique for a given choice of δ, additional structure is needed.

[9] For simplicity, I have drawn Figure 2 under the assumption that x_t is constant over time. Because x_t is constant, the expected collusive profit earned in each future period by a firm is the same as the profit it earns in the current period from colluding. Thus, to find the constrained sets for given realizations x_t, one needs to only evaluate the constraints at all possible combinations of prices for the two firms. However, when x_t is stochastic, numerically solving for the constraint sets for a given realization of x_t becomes more challenging as the constraints depend not only on the profits earned by the firms given the realization of x_t, but also on their expected profits over x. This leads to the curse of dimensionality, as one needs to evaluate the constraints at all possible combinations of prices for the two firms across all possible values of x. However, allowing x_t to be stochastic across periods would not fundamentally affect Figure 2. In particular, the feasible payoffs for a given value of δ and realizations of x_t would continue to form a convex set which would expand as δ increased.

Figure 2

Sustainable Payoffs In The Pricing Stage As A Function Of δ



2.2.2 Pareto Refinement

Given that there still exists a set of feasible subgame perfect Nash equilibria, I have to select which collusive equilibrium the firms choose. I assume that when firms collude, they choose actions for which it is impossible to make one firm better off without hurting its rival. Thus, the collusive equilibrium is Pareto optimal from the perspective of the firms. Because the demand shocks are exogenous, it is possible to separately solve for a collusive equilibrium at each possible realization of x_t. Given the realization, when profit possibility sets are convex, the collusive equilibrium in each period t can be represented as the solution to maximization problem faced by a third party coordinator, akin to a social planner.[10] The coordinator is appointed by the firms to maintain their collusive arrangement. To do so, the coordinator selects collusive actions $(d_{1t}^C, d_{2t}^C, p_{1t}^C, p_{2t}^C)$ to maximize its own welfare function, a weighted sum of the profits earned by these firms in period t. However, the coordinator faces the constraint that the collusive actions in each stage are incentive compatible for each firm given realizations of x_t and the firms' discount factor δ.

In each period, the coordinator's problem can be represented as a two-stage problem in which it selects an equilibrium by backwards induction. In the price setting stage, the coordinator chooses the following collusive prices to maximize its own payoffs subject to the incentive constraints for

[10] See Mas-Colell, Whinston and Green (1995), pages 558-566, for a complete discussion of representing a Pareto optimal equilibrium as the solution to the social planner's problem.

each of the firms that jointly define $\mathcal{P}_t$:

Def: The collusive prices $\left(p^C_{1t}(d_{1t}, d_{2t}; x_t, \delta, T), p^C_{2t}(d_{1t}, d_{2t}; x_t, \delta, T)\right)$ solve:

$$
\begin{aligned}
&\max_{p_{1t},p_{2t}} \sum_{b=1}^{2} \omega_b \pi_{bt}(d_{1t}, d_{2t}, p_{1t}, p_{2t}; x_t) \qquad (18)\\
&\text{st } \pi_{1t}(d_{1t}, d_{2t}, p_{1t}, p_{2t}; x_t) + \sum_{\tau=1}^{T} \delta^\tau \mathbb{E}_t \pi_{1t+\tau}(d_{1t+\tau}, d_{2t+\tau}, p^C_1(d_{1t+\tau}, d_{2t+\tau}; x_{t+\tau}), p^C_2(d_{1t+\tau}, d_{2t+\tau}; x_{t+\tau}); x_{t+\tau})\\
&\qquad \geq \pi^{C,D}_{1t}(d_{1t}, d_{2t}, p_{2t}; x_t) + \sum_{\tau=1}^{T} \delta^\tau \mathbb{E}_t \pi^{NE}_{1t+\tau}(x_{t+\tau})\\
&\quad \pi_{2t}(d_{1t}, d_{2t}, p_{1t}, p_{2t}; x_t) + \sum_{\tau=1}^{T} \delta^\tau \mathbb{E}_t \pi_{2t+\tau}(d_{1t+\tau}, d_{2t+\tau}, p^C_1(d_{1t+\tau}, d_{2t+\tau}; x_{t+\tau}), p^C_2(d_{1t+\tau}, d_{2t+\tau}; x_{t+\tau}); x_{t+\tau})\\
&\qquad \geq \pi^{C,D}_{2t}(d_{1t}, d_{2t}, p_{1t}; x_t) + \sum_{\tau=1}^{T} \delta^\tau \mathbb{E}_t \pi^{NE}_{2t+\tau}(x_{t+\tau})
\end{aligned}
$$

Then given these collusive prices, the coordinator chooses the following collusive products subject to the incentive constraints for each of the firms that jointly define $\mathcal{D}_t$:

Def: The collusive products $\left(d^C_{1t}(x_t, \delta, T), d^C_{2t}(x_t, \delta, T)\right)$ solve:

$$
\begin{aligned}
&\max_{d_{1t},d_{2t}} \sum_{b=1}^{2} \omega_b \pi_{bt}(d_{1t}, d_{2t}, p^C_{1t}(d_{1t}, d_{2t}; x_t), p^C_{2t}(d_{1t}, d_{2t}; x_t); x_t) \qquad (19)\\
&\text{st } \pi_{1t}(d_{1t}, d_{2t}, p^C_{1t}(d_{1t}, d_{2t}; x_t), p^C_{2t}(d_{1t}, d_{2t}; x_t); x_t) + \sum_{\tau=1}^{T} \delta^\tau \mathbb{E}_t \pi^{C,C}_{1t+\tau}(x_{t+\tau})\\
&\qquad \geq \pi^{D,C}_{1t}(d_{2t}; x_t) + \sum_{\tau=1}^{T} \delta^\tau \mathbb{E}_t \pi^{NE}_{1t+\tau}(x_{t+\tau})\\
&\quad \pi_{2t}(d_{1t}, d_{2t}, p^C_{1t}(d_{1t}, d_{2t}; x_t), p^C_{2t}(d_{1t}, d_{2t}; x_t); x_t) + \sum_{\tau=1}^{T} \delta^\tau \mathbb{E}_t \pi^{C,C}_{2t+\tau}(x_{t+\tau})\\
&\qquad \geq \pi^{D,C}_{2t}(d_{1t}; x_t) + \sum_{\tau=1}^{T} \delta^\tau \mathbb{E}_t \pi^{NE}_{2t+\tau}(x_{t+\tau})
\end{aligned}
$$

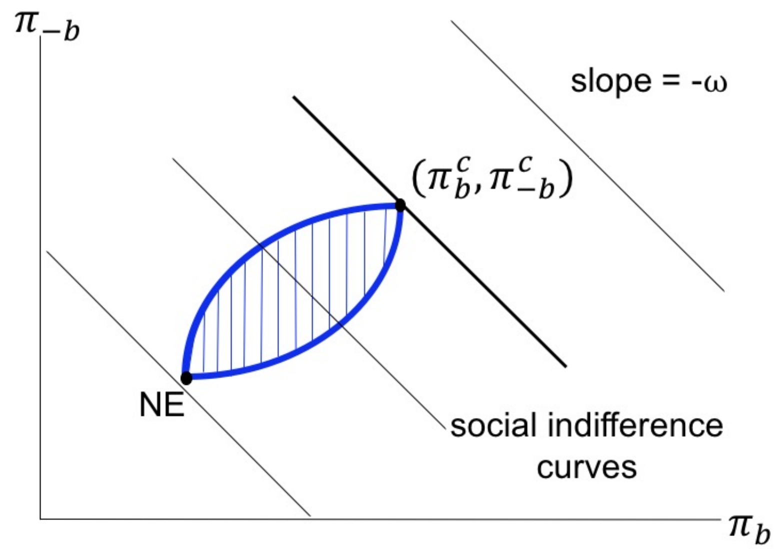
Thus, the refined equilibrium in this game is given by products (d^C_{1t}, d^C_{2t}) and prices (p^C_{1t}, p^C_{2t}) as defined in (18) and (19) and each firm b earns collusive profits $\pi^{CC}_{bt} = \pi_{bt}(d^C_{bt}, d^C_{-bt}, p^C_{bt}, p^C_{-bt})$ each period.

The coordinator's decision problem can be illustrated graphically in profit space. In Figure 3, I consider the pricing decision in (18).[11] The two constraints and the profit possibility frontier define

[11] Again, and without loss of generality, the incentive constraints are drawn under the assumption that x_t is constant over time for the reasons mentioned in footnote 9.

the lens of feasible payoffs available to the two firms for a given first stage product choice. Given that the coordinator places weights ω_1 and ω_2 on the profits earned by firm 1 and 2 respectively, the coordinator has linear indifference curves with slope $-\omega = \frac{\omega_1}{\omega_2}$. The coordinator chooses the collusive equilibrium at the point of tangency between its highest indifference curve and the constraints on payoffs induced by $\mathcal{P}_t$. Thus, the Pareto refinement ensures that the collusive equilibrium will lie on the frontier of the feasible payoff sets in each stage.[12]

Figure 3
Planner's Pricing Decision Problem



2.2.3 Deriving The Equivalence

I now show that with these refinements, the equilibrium selected by the coordinator is equivalent to the equilibrium in the reduced-form game. The first step is to rewrite the coordinator's problem as the decision problem made by the firms. Specifically, it is possible to derive an equivalent problem to (18) and (19) in which one firm b chooses actions for *both* itself and its rival in order to maximize its own payoff. In order to maintain the equivalence, an additional constraint needs to be imposed: the maximizing firm has to guarantee that its rival will receive a payoff at least as large as π^{CC}_{-bt}. (Mas-Colell, Whinston and Green (1995); pg. 562-566). Thus, the coordinator's problem can be rewritten in the following manner. In the second stage, firm b chooses prices at every possible product choice according to:

[12] Though not apparent in Figure 3, the equilibrium can depend on ω, especially as the set of feasible profits reaches the profit possibility frontier.

Def: The collusive equilibrium $\left(p_{1t}^C(d_{1t},d_{2t};x_t,\delta,T),p_{2t}^C(d_{1t},d_{2t};x_t,\delta,T)\right)$ simultaneously solves:

$$\max_{p_{bt},p_{-bt}} \ \pi_{bt}(d_{bt},d_{-bt},p_{bt},p_{-bt};x_t) \tag{20}$$

$$\text{st} \ \ \pi_{-bt}(d_{1t},d_{2t},p_{1t},p_{2t};x_t) \geq \pi_{-bt}(d_{bt},d_{-bt},p_{bt}^C(d_{bt},d_{-bt};x_t),p_{-bt}^C(d_{bt},d_{-bt};x_t);x_t)$$

$$\pi_{1t}(d_{1t},d_{2t},p_{1t},p_{2t};x_t)+\sum_{\tau=1}^{T}\delta^\tau \mathbb{E}_t\pi_{1t+\tau}(d_{1t+\tau},d_{2t+\tau},p_1^C(d_{1t+\tau},d_{2t+\tau};x_{t+\tau}),p_2^C(d_{1t+\tau},d_{2t+\tau};x_{t+\tau});x_{t+\tau})$$

$$\geq \pi_{1t}^{C,D}(d_{1t},d_{2t},p_{2t};x_t)+\sum_{\tau=1}^{T}\delta^\tau \mathbb{E}_t\pi_{1t+\tau}^{NE}(x_{t+\tau})$$

$$\pi_{2t}(d_{1t},d_{2t},p_{1t},p_{2t};x_t)+\sum_{\tau=1}^{T}\delta^\tau \mathbb{E}_t\pi_{2t+\tau}(d_{1t+\tau},d_{2t+\tau},p_1^C(d_{1t+\tau},d_{2t+\tau};x_{t+\tau}),p_2^C(d_{1t+\tau},d_{2t+\tau};x_{t+\tau});x_{t+\tau})$$

$$\geq \pi_{2t}^{C,D}(d_{1t},d_{2t},p_{1t};x_t)+\sum_{\tau=1}^{T}\delta^\tau \mathbb{E}_t\pi_{2t+\tau}^{NE}(x_{t+\tau})$$

Then given the optimal collusive prices associated with each choice of products, firm b chooses products in order to solve:

Def: The collusive equilibrium $\left(d_{1t}^C(x_t,\delta,T),d_{2t}^C(x_t,\delta,T)\right)$ simultaneously solves:

$$\max_{d_{bt}} \ \pi_{bt}(d_{bt},d_{-bt},p_{bt}^C(d_{1t},d_{2t};x_t),p_{-bt}^C(d_{1t},d_{2t};x_t);x_t) \tag{21}$$

$$\text{st} \quad \pi_{-bt}(d_{1t},d_{2t},p_{1t}^C(d_{1t},d_{2t};x_t),p_{2t}^C(d_{1t},d_{2t};x_t);x_t) \geq \pi_{-bt}^{C,C}(x_t)$$

$$\pi_{1t}(d_{1t},d_{2t},p_{1t}^C(d_{1t},d_{2t};x_t),p_{2t}^C(d_{1t},d_{2t};x_t);x_t)+\sum_{\tau=1}^{T}\delta^\tau \mathbb{E}_t\pi_{1t+\tau}^{C,C}(x_{t+\tau})$$

$$\geq \pi_{1t}^{D,C}(d_{2t};x_t)+\sum_{\tau=1}^{T}\delta^\tau \mathbb{E}_t\pi_{1t+\tau}^{NE}(x_{t+\tau})$$

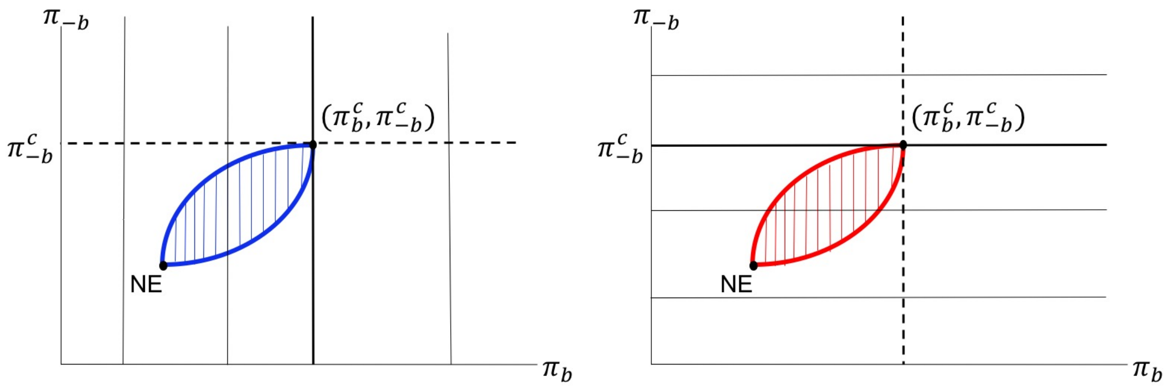
$$\pi_{2t}(d_{1t},d_{2t},p_{1t}^C(d_{1t},d_{2t};x_t),p_{2t}^C(d_{1t},d_{2t};x_t);x_t)+\sum_{\tau=1}^{T}\delta^\tau \mathbb{E}_t\pi_{2t+\tau}^{C,C}(x_{t+\tau})$$

$$\geq \pi_{2t}^{D,C}(d_{1t};x_t)+\sum_{\tau=1}^{T}\delta^\tau \mathbb{E}_t\pi_{2t+\tau}^{NE}(x_{t+\tau})$$

The identity of the firm making the decision does not matter here. If the decision is made by firm 1, the Lagrange multiplier on the first constraint in (20) and (21) is defined as ω_2/ω_1. Likewise, if firm 2 solves the problem, the Lagrange multiplier on the first constraint in (20) and (21) is given as ω_1/ω_2. The equivalence is illustrated in Figure 4. In the left panel, firm b chooses the actions to maximize its own payoff, guaranteeing at least the collusive equilibrium payoff to its rival. Thus, it has vertical indifference curves and faces an additional horizontal constraint. In the right panel, firm $-b$ chooses actions to maximize its own payoff, guaranteeing at least the collusive equilibrium to its rival. Thus, it has horizontal indifference curves and faces an additional vertical constraint. The equilibrium in both cases is the same as that depicted in Figure 3.

FIGURE 4

EQUIVALENCE OF EQUILIBRIA DEPENDING ON WHICH FIRM CHOOSES ACTIONS



In (20) and (21) the collusive equilibrium is represented as the solution to constrained maximization problems faced by *one* firm choosing the actions played by *both* firms. A key insight of BCLW (2004) is that the collusive equilibrium can also be expressed as the Nash equilibrium in a game where *every* firm, taking the choices of its rivals as given, independently and simultaneously chooses its own actions to maximize its profit subject to a set of constraints. These include the guarantee constraint: firm b must choose actions ensuring that its rival earns a profit in period t at least as large as that rival's collusive payoff from the coordinator given the actions of firm $-b$. In addition, firm b must choose actions in period t which satisfy its own incentive compatibility constraint given the actions of its rival. Specifically, the collusive equilibrium can be defined as follows:

Def: For all $b = \{1, 2\}$, the collusive equilibrium prices $(p^C_{1t}(d_{1t}, d_{2t}; x_t, \delta, T), p^C_{2t}(d_{1t}, d_{2t}; x_t, \delta, T))$ simultaneously solve:

$$V_{bt}(d_{bt}, d_{-bt}, p_{-bt}; x_t, \delta, T) = \max_{p_{bt}} \ \pi_{bt}(d_{bt}, d_{-bt}, p_{bt}, p_{-bt}; x_t) \tag{22}$$

$$\text{st } \pi_{-bt}(d_{1t}, d_{2t}, p_{1t}, p_{2t}; x_t) \geq \pi_{-bt}(d_{bt}, d_{-bt}, p^C_{bt}(d_{bt}, d_{-bt}, x_t), p^C_{-bt}(d_{bt}, d_{-bt}, x_t); x_t)$$

$$\pi_{bt}(d_{1t}, d_{2t}, p_{1t}, p_{2t}; x_t) + \sum_{\tau=1}^{T} \delta^\tau \mathbb{E}_t \pi_{bt+\tau}(d_{1t+\tau}, d_{2t+\tau}, p^C_1(d_{1t+\tau}, d_{2t+\tau}; x_{t+\tau}), p^C_2(d_{1t+\tau}, d_{2t+\tau}; x_{t+\tau}); x_{t+\tau})$$

$$\geq \pi^{C,D}_{bt}(d_{1t}, d_{2t}, p_{-bt}; x_t) + \sum_{\tau=1}^{T} \delta^\tau \mathbb{E}_t \pi^{NE}_{bt+\tau}(x_{t+\tau})$$

where:

$$V_{bt}(d_{bt}, d_{-bt}, p_{-bt}; x_t, \delta, T) = \pi_{bt}(d_{bt}, d_{-bt}, p^C_{bt}(d_{bt}, d_{-bt}, x_t), p^C_{-bt}(d_{bt}, d_{-bt}, x_t); x_t)$$

$$\& \ V_{-bt}(d_{bt}, d_{-bt}, p_{bt}; x_t, \delta, T) = \pi_{-bt}(d_{bt}, d_{-bt}, p^C_{bt}(d_{bt}, d_{-bt}, x_t), p^C_{-bt}(d_{bt}, d_{-bt}, x_t); x_t)$$

Then given the collusive prices, the collusive products can be defined as follows:

Def: For all $b = \{1, 2\}$, the collusive equilibrium products $\Big(d^C_{1t}(x_t, \delta, T), d^C_{2t}(x_t, \delta, T)\Big)$ simultaneously solve:

$$V_{bt}(d_{-bt}; x_t, \delta, T) = \max_{d_{bt}} \ \pi_{bt}(d_{bt}, d_{-bt}, p^C_{bt}(d_{1t}, d_{2t}; x_t), p^C_{-bt}(d_{1t}, d_{2t}; x_t); x_t) \tag{23}$$

$$\text{st} \qquad \pi_{-bt}(d_{1t}, d_{2t}, p^C_{1t}(d_{1t}, d_{2t}; x_t), p^C_{2t}(d_{1t}, d_{2t}; x_t); x_t) \geq \pi^{C,C}_{-bt}(x_t)$$

$$\pi_{bt}(d_{1t}, d_{2t}, p^C_{1t}(d_{1t}, d_{2t}; x_t), p^C_{2t}(d_{1t}, d_{2t}; x_t); x_t) + \sum_{\tau=1}^{T} \delta^\tau \mathbb{E}_t \pi^{C,C}_{bt+\tau}(x_{t+\tau})$$

$$\geq \pi^{D,C}_{bt}(d_{-bt}; x_t) + \sum_{\tau=1}^{T} \delta^\tau \mathbb{E}_t \pi^{NE}_{bt+\tau}(x_{t+\tau})$$

where:

$$V_{bt}(d_{-bt}; x_t, \delta, T) = \pi^{CC}_{bt}$$

$$\& \ V_{-bt}(d_{bt}; x_t, \delta, T) = \pi^{CC}_{-bt}$$

$$\& \ (\pi^{CC}_{1t}, \pi^{CC}_{2t}) = \text{the profits obtained at the coordinator's solution given } x_t, \delta, T$$

It can quickly be verified that the vector of actions which satisfy the system of first order conditions in the game just defined is the vector of collusive actions that solves the coordinator's problem. Notice that the guarantee constraints in the above game will be satisfied with equality for each firm in each stage, as neither firm will want to give to its rival a higher level of profit than it is constrained to offer. Therefore, each firm, given the actions of its rival, is constrained to choose actions guaranteeing that its rival will receive the collusive payoffs offered by the coordinator. As such, any vector of actions that satisfies the system of first order conditions will generate the same collusive profits for each firm as those obtained from the coordinator. In addition, each firm b chooses actions which satisfy its own incentive compatibility constraint given the actions played by its rival. Therefore, any vector of actions which simultaneously solves the system of first order conditions must satisfy the incentive constraints for all firms. One particular vector of actions which satisfies the incentive constraints is the vector of collusive actions which solves the coordinator's problem. And because the firms are constrained to obtain the collusive payoffs obtained from the coordinator, the vector of collusive actions chosen by the coordinator will solve the system of first order conditions to the above game for the firms. Therefore, the Pareto-refined subgame perfect Nash equilibrium of the repeated game in (19) and (18) can be expressed as the subgame perfect Nash equilibrium of a static game played by the firms with constraints in (22) and (23).

The first order condition for firm b in (22) can be written as:

$$\frac{\partial \pi_{bt}}{\partial p_{bt}} + \frac{\omega_{-b}/\omega_b}{1 + \lambda_{bt}(x_t, \delta, T)} \frac{\partial \pi_{-bt}}{\partial p_{bt}} = 0 \tag{24}$$

where ω_{-b}/ω_b is the Lagrange multiplier on the guarantee constraint for firm b and λ_{bt} is the

Lagrange multiplier on the guarantee constraint for firm b in (22). Notice that λ_{bt} is a function of x_t, δ, and T, and implicitly, the profit functions. The first order condition for firm b in (23) can be written as:

$$\frac{\mathbf{d}\pi_{bt}}{\mathbf{d}d_{bt}} + \frac{\omega_{-b}/\omega_b}{1+\mu_{bt}(x_t,\delta,T)}\frac{\mathbf{d}\pi_{-bt}}{\mathbf{d}d_{bt}} = 0 \tag{25}$$

where ω_{-b}/ω_b is the Lagrange multiplier on the guarantee constraint for firm b and μ_{bt} is the Lagrange multiplier on the incentive constraint for firm b in (23). Notice that μ_{bt} is a function of x_t, δ, and T, and implicitly, the profit functions.

The above analysis shows that the collusive outcome is equivalent to the outcome given by the SPNE of a one-shot game with reduced-form parameters $(\theta_{1t}, \theta_{2t})$ under the following four assumptions: 1. collusion involves Nash reversion as punishments, 2. the punishment lengths are given by T and ∞ for deviations in one or both stages respectively, 3. the collusive equilibrium is stationary, and 4. the collusive equilibrium is Pareto optimal where ω indexes the equilibrium choice on the frontier of feasible payoff set. Specifically, given T and ω, a mapping from the structural parameters, δ and the profit functions, to the reduced form parameters exists for each firm b in each period t and is given by:[13]

$$(\theta_{1,bt}, \theta_{2,bt}) = \left(\frac{\omega_{-b}/\omega_b}{1+\mu_{bt}(x_t,\delta,T)}, \frac{\omega_{-b}/\omega_b}{1+\lambda_{bt}(x_t,\delta,T)}\right) \tag{26}$$

2.2.4 Discussion Of The Reduced-Form Parameters

The mapping in (26) provides a theoretical justification for including different reduced-form parameters in the product stage and pricing stage. In particular, the price and product collusion parameters differ because $\theta_{1,bt}$ is a function of μ_{bt} while $\theta_{2,bt}$ is a function of λ_{bt}. μ_{bt} measures the marginal benefit of relaxing the incentive constraint for firm b on the set of products for which collusion is feasible while λ_{bt} measures the marginal benefit of relaxing the incentive constraint for firm b on the set of feasible prices. These sets depend on the profitability of deviation in the product and pricing stages respectively. Because the profitability of deviation is likely to differ across the stages, so too should the multipliers and therefore the reduced-form parameters measuring collusion.

In order to validate the mapping, it is important to verify that it generates profit weights conforming to the predictions from theory at extreme values of δ. When $\delta = 0$, the firms play the one-shot subgame perfect Nash equilibrium, implying that the profit weights for each firm in each stage should be 0. This can be verified by rewriting (24) as $\frac{1+\lambda_{bt}}{\omega_{-b}/\omega_b} = -\frac{\partial\pi_{-bt}/\partial p_{bt}}{\partial\pi_{bt}/\partial p_{bt}}$. At the subgame perfect Nash equilibrium, $\frac{\partial\pi_{bt}}{\partial p_{bt}} = 0$, implying that $\lambda_{bt} \to \infty$ as $\delta \to 0$. The same analysis would show that $\mu_{bt} \to \infty$ as $\delta \to 0$. Likewise, when δ is large and the collusive equilibria lie on the profit possibility frontier, the incentive constraints in each stage are no longer binding. Thus, the λ's and μ's $\to 0$ as $\delta \to 1$ and collusion parameters in each stage become ω_{-b}/ω_b for firm b.

[13] The mapping is not one-to-one with respect to δ. In particular, there exists a lower bound $\underline{\delta}$ such that, for all $\delta > \underline{\delta}$, the equilibrium lies on the profit possibility frontier.

Finally, as will become apparent in Section 4.2.2, it is important to understand how the reduced-form collusion parameters in the pricing stage would change if either of the firms deviated in the product stage. Recall that following a deviation in the product stage, the firms can still choose to collude in the pricing stage. If each firm plays the collusive price specified for the offered set of products, then the firms are punished for T periods. If either firm plays different prices, the punishment lasts forever. I have assumed that the firms are patient enough such that they will not deviate in both stages. Therefore, the price collusion parameters will not be zero following a deviation in the product stage. However, if the firms are patient enough, any possible menu of off path prices is feasible, suggesting that the reduced-form pricing parameters could take on any value. Notice that, for every choice of products, the firms could choose off path prices such that the value of the reduced-form pricing parameter implied by these prices is the same as it is on path. Since this is one possible equilibrium, I assume it is the one that the firms play. Therefore, the price collusion parameters, θ_2, are the same at each possible product deviation as they are on path.

2.3 Extension To Multiple Markets

When estimating reduced-form collusion parameters, it is important to consider whether the parameters are robust to the critique raised in Corts (1999). If firms compete in one market, then a researcher must hold the reduced-form collusion parameters fixed over time and observe multiple time periods in order to identify those parameters. However, in the structural model, the collusive equilibria in each period will depend on the sets of products and prices that were feasible for the firms to offer. These sets $\mathcal{D}_t$ and $\mathcal{P}_t$ will vary over time with changes in market characteristics. This will lead to changes in the Lagrange multipliers (λ's and μ's) implying that the true values of θ_{1t} and θ_{2t} are not fixed over time.

In my empirical setting, Ben & Jerry's and Häagen-Dazs compete in many geographic markets each period. With this cross-sectional variation, it is possible to estimate a different collusion parameter each period. At first glance, this seems to replace one problem with another: in order to identify conduct with cross-sectional data, instruments that shift and rotate the demand curve cross-sectionally are needed to identify conduct. But, if the true conduct parameters vary over time with demand shocks, won't they now also vary across markets within a period? However, Sullivan (2017) shows that, under an additional equilibrium refinement, the Lagrange multipliers are constant across markets within a period.[14] The ability of firms to compete in multiple markets each period expands the set of strategies available to the firms in the repeated game, even within the class of stick and carrot punishments. For instance, if a firm were to deviate from the collusive arrangement in only market m during period t, do the other firms punish that defection in just market m or in a larger set of markets? Additionally then, I assume that if a firm deviates in *any* market, it is punished by reversion to the one-shot subgame perfect Nash equilibrium in *all* markets. In a model with perfect information, punishing a deviating firm in all markets as opposed to a subset

[14] Sullivan (2017) considers the infinitely repeated quantity setting game from Corts (1999) but the result extends to this game as well.

of markets serves as the harshest punishment in this class and supports the highest collusive payoffs. Thus, conditional on the use of stick and carrot punishments, Nash reversion in all markets is the most efficient punishment scheme. Because punishment occurs in all markets, a firm that deviates will do so in all markets. Thus, the benefit of deviating in each stage in *any* market depends on the sum of the deviation profits across *all* markets. This ensures that the marginal benefit of deviating in any market is the same. Thus, the Lagrange multipliers on the constraints induced by $\mathcal{D}_t$ and $\mathcal{P}_t$ are constant across markets in a period and the reduced-form parameters implied by the structural model will not vary cross-sectionally.

Formally, let markets be indexed by $m = 1, \ldots, M$. In each market, firms observe realizations of random demand shocks before choosing actions in that market. Therefore, let $\mathbf{x}_t = \{x_{t1}, \ldots, x_{tM}\}$. Furthermore, let $\mathbf{d}_{bt} = \{d_{bt1}, \ldots, d_{btM}\}$ and $\mathbf{p}_{bt} = \{p_{bt1}, \ldots, p_{btM}\}$ be vectors of the product choices and prices for each firm in each market. The feasible sets of actions (16) and (17) can now be written as:

$$\begin{aligned}\mathcal{P}_t(\mathbf{d}_{1t},\mathbf{d}_{2t};\mathbf{x}_t,\delta,T) = \Bigg\{(\mathbf{p}_{1t},\mathbf{p}_{2t}) \Bigg| & \sum_{m=1}^{M} \pi_{btm}(d_{1tm}, d_{2tm}, p_{1tm}, p_{2tm}; x_{tm}) \\ & + \sum_{m=1}^{M}\sum_{\tau=1}^{T} \delta^{\tau}\mathbb{E}_t \pi_{bt+\tau m}(d_{1t+\tau m}, d_{2t+\tau m}, p_1^C(d_{1t+\tau m}, d_{2t+\tau m}; x_{t+\tau m}), p_2^C(d_{1t+\tau m}, d_{2t+\tau m}; x_{t+\tau m}); x_{t+\tau m}) \\ & \geq \sum_{m=1}^{M} \pi_{btm}^{C,D}(d_{1tm}, d_{2tm}, p_{-btm}; x_{tm}) + \sum_{m=1}^{M}\sum_{\tau=1}^{T} \delta^{\tau}\mathbb{E}_t \pi_{bt+\tau m}^{NE}(x_{t+\tau m}) \qquad \forall\, b = 1,2\Bigg\}\end{aligned} \tag{27}$$

and

$$\begin{aligned}\mathcal{D}_t(x_t,\delta,T) = \Bigg\{(d_{1t},d_{2t}) \Bigg| & \sum_{m=1}^{M} \pi_{btm}(d_{1tm}, d_{2t}, p_{1tm}^C(d_{1tm}, d_{2tm}; x_{tm}), p_{2tm}^C(d_{1tm}, d_{2tm}; x_{tm}); x_{tm}) \\ & + \sum_{m=1}^{M}\sum_{\tau=1}^{T} \delta^{\tau}\mathbb{E}_t \pi_{bt+\tau m}(d_{1t+\tau m}^C, d_{2t+\tau m}^C, p_1^C(d_{1t+\tau m}^C, d_{2t+\tau m}^C; x_{t+\tau m}), p_2^C(d_{1t+\tau m}^C, d_{2t+\tau m}^C; x_{t+\tau m}); x_{t+\tau m}) \\ & \geq \sum_{m=1}^{M} \pi_{btm}^{D,C}(d_{-btm}; x_{tm}) + \sum_{m=1}^{M}\sum_{\tau=1}^{T} \delta^{\tau}\mathbb{E}_t \pi_{bt+\tau m}^{NE}(x_{t+\tau m}) \qquad \forall\, b = 1,2\Bigg\}\end{aligned} \tag{28}$$

With the feasibility constraints so defined, the multi-market analog to (22) and (23) can be written as:

Def: For all $b = \{1, 2\}$, the collusive equilibrium prices $(\mathbf{p}^C_{1t}(\mathbf{d}_{1t}, \mathbf{d}_{2t}; \mathbf{x}_t, \delta, T), \mathbf{p}^C_{2t}(\mathbf{d}_{1t}, \mathbf{d}_{2t}; \mathbf{x}_t, \delta, T))$ simultaneously solve:

$$V_{bt}(\mathbf{d}_{bt}, \mathbf{d}_{-bt}, \mathbf{p}_{-bt}; \mathbf{x}_t, \delta, T) = \max_{p_{bt1},\dots,p_{btM}} \sum_{m=1}^{M} \pi_{btm}(d_{btm}, d_{-btm}, p_{btm}, p_{-btm}; x_{tm})$$

$$\begin{aligned}
\text{st } & \sum_{m=1}^{M} \pi_{-btm}(d_{1tm}, d_{2tm}, p_{1tm}, p_{2tm}; x_{tm}) \\
& \quad \geq \sum_{m=1}^{M} \pi_{-btm}(d_{btm}, d_{-btm}, p^C_{btm}(d_{btm}, d_{-btm}, x_{tm}), p^C_{-btm}(d_{btm}, d_{-btm}, x_{tm}); x_{tm}) \\
& \sum_{m=1}^{M} \pi_{btm}(d_{1tm}, d_{2tm}, p_{1tm}, p_{2tm}; x_{tm}) \\
& \quad + \sum_{m=1}^{M} \sum_{\tau=1}^{T} \delta^\tau \mathbb{E}_t \pi_{bt+\tau m}(d_{1t+\tau m}, d_{2t+\tau m}, p^C_1(d_{1t+\tau m}, d_{2t+\tau m}; x_{t+\tau m}), p^C_2(d_{1t+\tau m}, d_{2t+\tau m}; x_{t+\tau m}); x_{t+\tau m}) \\
& \quad \geq \sum_{m=1}^{M} \pi^{C,D}_{btm}(d_{1tm}, d_{2tm}, p_{-btm}; x_{tm}) + \sum_{m=1}^{M} \sum_{\tau=1}^{T} \delta^\tau \mathbb{E}_t \pi^{NE}_{bt+\tau m}(x_{t+\tau m})
\end{aligned}$$

where:

$$V_{bt}(\mathbf{d}_{bt}, \mathbf{d}_{-bt}, \mathbf{p}_{-bt}; \mathbf{x}_t, \delta, T) = \sum_{m=1}^{M} \pi_{btm}(d_{btm}, d_{-btm}, p^C_{btm}(d_{btm}, d_{-btm}, x_{tm}), p^C_{-btm}(d_{btm}, d_{-btm}, x_{tm}); x_{tm})$$

$$\& \ V_{-bt}(\mathbf{d}_{bt}, \mathbf{d}_{-bt}, \mathbf{p}_{bt}; \mathbf{x}_t, \delta, T) = \sum_{m=1}^{M} \pi_{-btm}(d_{btm}, d_{-btm}, p^C_{btm}(d_{btm}, d_{-btm}, x_{tm}), p^C_{-btm}(d_{btm}, d_{-btm}, x_{tm}); x_{tm})$$

Then given the collusive prices, the collusive products can be defined as follows:

Def: For all $b = \{1, 2\}$, the collusive equilibrium products $\left(\mathbf{d}^C_{1t}(x_t, \delta, T), \mathbf{d}^C_{2t}(x_t, \delta, T)\right)$ simultaneously solve:

$$\tag{29}$$

$$V_{bt}(\mathbf{d}_{-bt}; \mathbf{x}_t, \delta, T) = \max_{d_{bt1},\dots,d_{btM}} \sum_{m=1}^{M} \pi_{btm}(d_{btm}, d_{-btm}, p^C_{btm}(d_{1tm}, d_{2tm}; x_{tm}), p^C_{-btm}(d_{1tm}, d_{2tm}; x_{tm}); x_{tm})$$

$$\begin{aligned}
\text{st } \quad & \sum_{m=1}^{M} \pi_{-btm}(d_{1tm}, d_{2tm}, p^C_{1tm}(d_{1tm}, d_{2tm}; x_{tm}), p^C_{2tm}(d_{1tm}, d_{2tm}; x_{tm}); x_{tm}) \geq \sum_{m=1}^{M} \pi^{C,C}_{-btm}(x_{tm}) \\
& \sum_{m=1}^{M} \pi_{btm}(d_{1tm}, d_{2tm}, p^C_{1tm}(d_{1tm}, d_{2tm}; x_{tm}), p^C_{2tm}(d_{1tm}, d_{2tm}; x_{tm}); x_{tm}) \\
& \quad + \sum_{m=1}^{M} \sum_{\tau=1}^{T} \delta^\tau \mathbb{E}_t \pi^{C,C}_{bt+\tau m}(x_{t+\tau m}) \geq \sum_{m=1}^{M} \pi^{D,C}_{btm}(d_{-btm}; x_{tm}) + \sum_{m=1}^{M} \sum_{\tau=1}^{T} \delta^\tau \mathbb{E}_t \pi^{NE}_{bt+\tau m}(x_{t+\tau m})
\end{aligned} \tag{30}$$

where:

$$V_{bt}(\mathbf{d}_{-bt};\mathbf{x}_t,\delta,T) = \sum_{m=1}^{M} \pi_{btm}^{CC}$$

$$\& \ V_{-bt}(\mathbf{d}_{bt};\mathbf{x}_t,\delta,T) = \sum_{m=1}^{M} \pi_{-btm}^{CC}$$

$$\& \ (\pi_{1tm}^{CC}, \pi_{2tm}^{CC}) = \text{the profits obtained at the coordinator's solution given } \mathbf{x}_t, \delta, T$$

Notice that for a given period t, the reduced-form profit weights θ_{1t}, θ_{2t} will be constant across markets. The key insight is that in the single market model, the constraints in the optimization problems (22) and (23) are subject to change over time. leading to changes over time in the Lagrange multipliers λ_b and μ_b which lead to changes in θ_{1t} and θ_{2t}. Here, the constraints faced by firm b in choosing its actions in market m during period t are defined by the sum of payoffs across all the M markets in period t. Thus, the firm faces the same set of constraints in every market during period t. Because the constraints are not changing, the same Lagrange multiplier applies to the firm's decision in each market m during period t. Thus, the reduced-form collusion parameters do not vary with demand shocks across markets in a period.[15] Therefore, these shocks can be used as instruments that identify the reduced-form collusion parameters, and it is possible to estimate reduced-form collusion parameters in each period without running afoul of the Corts Critique.[16]

3 The Market For Super-Premium Ice Cream

3.1 Market Definitions

Relevant Product Market: For the purposes of this paper, the relevant product market consists of super-premium ice cream sold by the pint in supermarkets. The USDA defines four categories of ice cream: super-premium, premium, standard, and economy. Super-premium ice cream is distinguished as "tend(ing) to have very low overrun (air in the ice cream) and high fat content, and the manufacturer uses the best quality ingredients."[17] There is considerable precedent for defining super-premium ice cream as a distinct product market: the economics literature,[18] antitrust decisions,[19] and trade publications have all used this assumption.

[15] While now robust to the Corts critique, these parameter values are still sensitive to the Lucas Critique. Under general counterfactual settings, the value of the collusive parameters is not necessarily fixed. As will be discussed below, the policy relevant counterfactual I consider is one in which the firms cannot collude, and it is known that the reduced-form collusion parameters will be 0.

[16] In practice, I estimate one set of collusion parameters for 2013 using monthly panel data. In doing so, I implicitly assume that the price and product choices for each month were all made at the beginning of 2013. If deviation occurs during the year, it is not punished until the following year. While it is unlikely that firms punish on a weekly basis, it is also likely that punishment might occur with higher frequency. In the future, I plan on estimating different collusion parameters for each month to test the robustness of my estimates to the Corts Critique.

[17] http://www.ams.usda.gov/sites/default/files/media/CID%20Ice%20Cream%2C%20Sherbet%2C%20Fruit%20and%20Juice%20Bars%2C%20Ices%2C%20and%20Novelties.pdf

[18] DMS(2009) study the market for premium ice cream as distinct from super-premium.

[19] https://www.ftc.gov/sites/default/files/documents/cases/2003/06/dreyercomplaint.htm

There are three important exclusions being made here. First, I am assuming that the market for prepackaged ice cream to be consumed at home is distinct from the market for ice cream sold by the scoop in ice cream shops. Second, I am assuming that prepackaged pints sold in supermarkets exist in a separate market from prepackaged or hand packed tubs that a brand might sell at its scoop shops and pints sold in convenience stores and drug stores. These first two assumptions seem largely innocuous. The third assumption, that super-premium ice cream resides in a market distinct from premium, standard, and economy ice cream, is more readily contestable. However, there is a significant degree of horizontal and vertical differentiation between super-premium brands and their closest competitors, the premium brands. In addition to the differences in fat and air content and ingredient quality, super-premium brands are sold in "pints" while premium brands are sold in quarts, half gallons, and gallons. Super-premium and premium brands tend to be sold in different parts of the freezer section, reducing cross-category comparison and substitution.

Relevant Geographic Market: I define the relevant geographic market at the supermarket level. Consumers choose the stores at which they shop for groceries based on the store's location, the set of product bundles offered, and the price of those bundles. Given that the price for a weekly bundle of groceries for a family of four is typically in the hundreds of dollars, a small change in the price of ice cream has a negligible impact on the price of the bundle. Furthermore, consumers tend to shop for their entire bundle of groceries at one supermarket in part because of transportation costs and the costs of learning a store's layout. Thus, they are unlikely to know the prices charged for ice cream in another store. For these reasons, I surmise that the cross price elasticity between stores for ice cream may be very low.

3.2 Who Chooses Prices and Flavors?

When considering the supply side of the model, there are a variety of actors in the vertical stream of production of ice cream. In particular, there are manufacturers and supermarket retailers. In some instances there are independent distributors. This opens up the possibility for complex vertical arrangements and raises an important question: who chooses the flavors and sets the prices in each market? For the purposes of this paper, I ignore the possibility of vertical interactions and assume that Ben & Jerry's and Häagen-Dazs choose the flavors to offer in each market and set the prices for those products. I justify this assumption in two ways. First, it is the operating assumption in the literature. As stated in DMS (2009)

> "*Since our data is aggregated across stores in a market area, we consider the manufacturers' product-choice decisions of which flavors to offer at the market level abstracting from the manufacturer-retailer interaction. The institutional realities in the ice cream industry suggest that manufacturers have substantial control over the varieties placed in the supermarkets. Ice cream is not handled through supermarket warehouses but through a direct-to-store distribution network. Ice-cream manufacturers 'rent' freezer space in the stores and retain full responsibility for what to stock.*"

Though I am not aggregating my data, the validity of the above argument should not be affected by aggregation. For the assumption to be true at the city level, it must in some sense hold at every store in the city.

Secondly, the reliance on direct-to-store distribution (DSD) strengthens the credibility of this assumption. Direct-to-store distributors not only deliver the product to the supermarket, but are responsible for stocking the store's shelves with that product. According to a 2008 report from the Grocery Manufacturers Association, "Knowledgeable representatives of suppliers of DSD products are in stores multiple times a week merchandising products ... the supplier assumes the costs for delivery, inventory management and merchandising." Dari Farms, a direct-to-store distributor of ice cream touts amongst its services, "optimiz(ing) product mix and profitability." Given this, I feel comfortable assuming that the supermarkets are not choosing products and prices.

It is possible that distributors are choosing the prices and product offerings. This seems unlikely though. These distributors have been granted exclusive territories by the manufacturers. My conjecture is that the cost to manufacturers from switching distributors is low. Thus, it is likely that distributors have minimal bargaining power with manufacturers and would be loathe to contradict their wishes. For these reasons, I assume that the choices of products and prices are made by Ben & Jerry's and Häagen-Dazs.

4 Empirical Model

4.1 Consumer Demand

Consumers $i = 1, \ldots, M_t$ shop for groceries in market t, where $t = 1, \ldots, T$ indexes a unique supermarket-week pair. Each consumer can purchase one pint of ice cream from either Ben & Jerry's or Häagen-Dazs or consume the outside option.[20] The utility consumer i receives from purchasing flavor j produced by brand b in market t is given by:

$$U_{ijbt} = \beta_i X_{bj} + \alpha p_{bt} + \tilde{\xi}_{bjt} + \epsilon_{ijbt} \tag{31}$$

where X_{bj} is a vector of observable product characteristics and p_{bt} is the price charged by firm b for all the flavors it sells in market t.[21] Market specific tastes for each product are captured by $\tilde{\xi}_{bjt}$ while ϵ_{ijbt} is an consumer-specific idiosyncratic term. Consumer preferences for the characteristics in X are assumed normally distributed across the population such that $\beta_i \sim \mathcal{N}(\bar{\beta}, \sigma_\beta)$, while α, the coefficient on price, is constant across consumers.

The observable characteristics in X include a constant, a brand dummy, indicators for whether a given flavor has a vanilla, chocolate, coffee, or fruit base, and a categorical variable that measures

[20] I do not allow consumers to store ice cream for future consumption. Hendel and Nevo (2006*a*) and Hendel and Nevo (2006*b*) highlight the potential problems in doing so, especially when one considers weekly markets. However, due to the high storage costs both in terms the cost of freezer shelf space faced by retailers and the temptation costs faced by consumers, I do not feel storability is much of an issue in this setting.

[21] The assumption is consistent with DMS (2009) and much of the literature on pricing at grocery stores. However, it does preclude sales for a strict subset of flavors offered in a market.

the number of mix-ins in that flavor. For example, Chunky Monkey is banana ice cream with dark chocolate chunks and walnuts. Thus, it is represented as having a fruit base and two mix-ins. Meanwhile coffee ice cream has a coffee base and no mix-ins.

Admittedly, X excludes many observable flavor characteristics which are important to consumer utility. Diverse flavors like peanut butter cup and mint oreo are observationally equivalent in X. Also, consumer utility from a flavor should depend on the interaction between the base and the type of mix-ins used, which are not measured in X. For example, most consumers would prefer chocolate ice cream with brownies to chocolate ice cream with grapes. Because of computational and data limitations, I have limited X to the characteristics I think are most important in describing a flavor.[22]

To account for these limitations and improve the fit of the model, Nevo (2001) recommends the inclusion of fixed effects. As such, I decompose the market specific product tastes as follows:

$$\tilde{\xi}_{bjt} = \xi_{bj} + \xi_s + \xi_m + \xi_y + \xi_{bjt} \tag{32}$$

where ξ_{bj}, ξ_s, ξ_m, and ξ_y are respectively product, store, month, and year fixed effects.[23] ξ_{bjt} measures the deviation in market specific product tastes from these means. The inclusion of product fixed effects allows the mean utility for products with the same value of X to differ. Each product fixed effect also measures the average utility derived from specific base flavor and mix-in combinations. The month fixed effects are also very important since ice cream is a product that also exhibits a high degree of seasonality.[24]

I make two assumptions concerning the demand shocks ξ_{bjt}. The first is a timing assumption of the sort discussed in Ackerberg and Hahn (2015) and Ackerberg (2016). Let $\xi_{bt} = [\xi_{b1t}, \ldots, \xi_{bJt}]$ be the vector of demand shocks faced by brand b in market t. I first assume that each firm chooses both the products it offers and the price it charges in market t before the demand shocks ξ_{bt} and ξ_{-bt} are realized.[25] Second, I assume that the ξ_{bjt} are *iid* across all products within a market and across markets (cross-sectionally and over time).

In addition to the utility of the inside goods produced by Ben & Jerry's and Häagen-Dazs, I have to specify the utility received by consumers who do not purchase a flavor from either brand. I have chosen the traditional normalization: the utility consumer i receives from purchasing the outside option in market t is given by $u_{i0t} = \epsilon_{i0t}$.

I impose two additional assumptions which are standard in the literature. The first is that the

[22] To make the substitution patterns fully flexible, one would need to include dummies for all base flavors, all mix-ins, and their interactions. Unfortunately, this is infeasible. In particular, Ben & Jerry's alone uses 53 base flavors and 100 mix-ins in my sample.

[23] I discuss how $\bar{\beta}$ and ξ_{bj} are separately identified in Section 6.1.

[24] Though I do not allow for seasonal variation in the utility of individual flavors, seasonal flavors are only sold during certain months of the year. Thus, the product fixed effects are only measured during the months in which each flavor is sold. When constructing moment inequalities, I am careful not to consider deviations in which seasonal flavors are offered out of season.

[25] Admittedly, this assumption is particularly strong. In future work, I plan on relaxing it by allowing ξ_{bjt} to follow and AR(1) process as in Sweeting (2013). If $\xi_{bjt} = \rho\xi_{bjt-1} + \nu_{bjt}$, the demand parameters are still identified under the assumption that firms choose products and prices in market t before ν_{bjt} is realized.

idiosyncratic errors ϵ_{ijbt} and ϵ_{i0t} are *iid* draws from a Type I extreme value distribution. The second is that each consumer purchases one unit of the good that gives her the highest utility, including the outside good. Then, if each firm can choose a subset of J products to offer, the market share of product j produced by brand b in market t is represented as:

$$s_{bjt} = d_{bjt} \int \frac{\exp(\beta_i X_{bj} + \alpha p_{bt} + \xi_{bj} + \xi_s + \xi_y + \xi_m + \xi_{bjt})}{1 + \sum_{h \in \{b,-b\}} \sum_{k=1}^{J} d_{bkt} \exp(\beta_i X_{hk} + \alpha p_{ht} + \xi_{hk} + \xi_s + \xi_y + \xi_m + \xi_{hkt})} f(\beta_i) \mathrm{d}\beta_i \tag{33}$$

where d_{bjt} is an indicator for whether firm b offers product j in market t.

4.2 Supply

Following the theoretical model presented in Section 2, I model Ben & Jerry's and Häagen-Dazs as competing in static reduced-form game which has two stages: first, firms choose a subset of J products to offer in each market t, then they choose prices. It is unlikely that the firms make decisions each week, especially with respect to their product choices. Therefore, I assume that this game is played monthly and, on the supply side, market t corresponds to a supermarket-month.

I now specify the reduced-form game played by Ben & Jerry's and Häagen-Dazs under the timing and distributional assumptions concerning ξ_{bjt} that were made in the previous section. In the first stage, each firm b solves:

$$\max_{d_{b1},\ldots,d_{bT}} \sum_t \mathbb{E}_\xi[\pi_{bt}] + \theta_{1,b}\mathbb{E}_\xi[\pi_{-bt}] \tag{34}$$

where $d_{bt} = [d_{b1t}, \ldots, d_{bJt}]$. Meanwhile in the second stage, after observing the flavors offered by its rival, firm b chooses the prices to charge in each market in order to solve:

$$\max_{p_{b1},\ldots,p_{bT}|d_{b1},\ldots,d_{bT}} \sum_t \mathbb{E}_\xi[\pi_{bt}] + \theta_{2,b}\mathbb{E}_\xi[\pi_{-bt}] \tag{35}$$

where the profit function π_{bt} is defined as

$$\pi_{bt}(d_{bt}, d_{-bt}, p_{bt}, p_{-bt}) = M_t(p_{bt} - c_{bt}) \sum_{j=1}^{J} d_{bjt} s_{bjt} - R_{bt} \tag{36}$$

and M_t is the market size. The profit function includes two cost parameters. The first cost parameter, c_{bt}, represents the marginal costs of production, distribution, and retail. As with price, the marginal cost is assumed to be constant across all flavors sold by a brand in a given market.[26] I

[26] This is also the assumption made in DMS (2009), and is relatively innocuous given that all flavors are largely composed of similar amounts of cream, eggs, and sugar. Thus, the ingredients that distinguish flavors are likely a small part of the production cost. In addition, distribution costs, especially refrigerated transportation, are likely a major component of c and would be constant across flavors.

assume that c_{bt} is linear in a set cost shifters

$$c_{bt} = \mathrm{w}_{bt}\gamma + \omega_{bt} \tag{37}$$

where w_{bt} and ω_{bt} are observed and unobserved cost shifters respectively. The second cost parameter, R_{bt}, represents a fixed cost that brand b has to pay to the retailer in market t. I assume this cost is a linear function of N_{bt}, the number of flavors that firm b offers in market t.

$$R_{bt} = \eta_b N_{bt} \tag{38}$$

One interpretation of η_b is as the per-flavor rental cost of freezer shelf space. Freezer space in a supermarket is limited and supermarkets face an opportunity cost to stocking an additional flavor produced by brand b. While I do not model the bargaining process between the supermarkets and the brands, this interpretation allows the variable profits to be split between the manufacturers and the retailers. Alternatively, η_b could capture other costs that scale linearly with the number of varieties produced brand b including the cost of switching a production line from one flavor to another flavor.

With the profit function defined, the SPNE of this reduced-form game can be found by backwards induction. Thus, I now consider the firms' second stage pricing decisions.

4.2.1 Second Stage Pricing Decision

In the second stage, firms take the choice of products in the first stage as given. Thus in each market, firm b chooses its price p_{bt} in order to solve the following maximization problem.

$$\begin{aligned} \max_{p_{b1},\dots,p_{bT}} \sum_{t=1}^{T} \left(M_t(p_{bt} - c_{bt}) \sum_{j=1}^{J} d_{bjt}\mathbb{E}_\xi[s_{bjt}] - R_{bt} \right) \\ + \theta_{2,b} \sum_{t=1}^{T} \left(M_t(p_{-bt} - c_{-bt}) \sum_{j=1}^{J} d_{-bjt}\mathbb{E}_\xi[s_{-bjt}] - R_{-bt} \right) \end{aligned} \tag{39}$$

There are T first order conditions associated with this optimization problem. Because the markets are independent, the first order condition associated with p_{bt} depends only on variables specific to market t. Thus, each first order condition can be considered separately. Rearranging the first order condition governing firm b's pricing decision in market t leads to the following expression for the price cost markup:

$$p_{bt} - c_{bt} = \frac{\sum_{j=1}^{J} d_{bjt}\mathbb{E}_\xi[s_{bjt}] + \theta_{2,b}\Delta_{b,-b,t} \sum_{j=1}^{J} d_{-bjt}\mathbb{E}_\xi[s_{-bjt}]}{\sum_{j=1}^{J} d_{bjt}(-\frac{\partial \mathbb{E}_\xi[s_{bjt}]}{\partial p_{bt}}) - \theta_{2,b}\theta_{2,-b}\Delta_{b,-b}\sum_{j=1}^{J} d_{-bjt}\frac{\partial \mathbb{E}_\xi[s_{-bjt}]}{\partial p_{bt}}} \tag{40}$$

where

$$\Delta_{b,-b} = \frac{\sum_{j=1}^{J} d_{bjt} \frac{\partial \mathbb{E}_\xi[s_{bjt}]}{\partial p_{-bt}}}{\sum_{j=1}^{J} d_{-bjt}(-\frac{\partial \mathbb{E}_\xi[s_{-bjt}]}{\partial p_{-bt}})} \tag{41}$$

The functional form of the markup is fairly intuitive. When the price collusion parameters equal 0, (40) reduces to the ratio of firm b's total market share to its total own price elasticity, which is the standard markup in the literature given Nash-Bertrand competition and logit demand. As the collusion parameters increase above 0, the firms begin to internalize the impact of their pricing decisions on their rival's profit. Thus, the numerator of the markup is the weighted sum of firm b's total market share and the total market share of its rival. Likewise, the denominator is the weighted difference of firm b's own price elasticity and the cross price elasticity.

The expression in (40) also gives rise to two comparative statics that a reader would expect in a model of collusion. First, increasing the degree to which firms collude leads to higher prices and larger markups. An increase in either θ_b or θ_{-b} results in an increase in the markup, and because marginal cost is exogenous, an increase in the markup must be the result of an increase in price. Secondly, the effect of collusion on prices and markups depends on the degree to which firms compete in the market. All else equal, there is greater scope for collusion to increase prices in markets where firms offer direct substitutes than in markets where the products offered by firms are not perceived as substitutable. Cross brand substitutability is captured by $\Delta_{b,-b,t}$. Given a marginal increase in p_{-bt}, a subset of consumers will stop purchasing products made by firm $-b$. $\Delta_{b,-b,t}$ measures the fraction of those customers who switch to purchasing products produced by firm b as opposed to those that switch to consuming the outside option. In (40) the derivative of the markup with respect to either collusion parameter is increasing in $\Delta_{b,-b,t}$

Combining the first order condition defined in (40) with the marginal cost specification (37) yields the following equation.

$$p_{bt} - \frac{\sum_{j=1}^{J} d_{bjt}\mathbb{E}_\xi[s_{bjt}] + \theta_{2,b}\Delta_{b,-b}\sum_{j=1}^{J} d_{-bjt}\mathbb{E}_\xi[s_{-bjt}]}{\sum_{j=1}^{J} d_{bjt}(-\frac{\partial \mathbb{E}_\xi[s_{bjt}]}{\partial p_{bt}}) - \theta_{2,b}\theta_{2,-b}\Delta_{b,-b}\sum_{j=1}^{J} d_{-bjt}\frac{\partial \mathbb{E}_\xi[s_{-bjt}]}{\partial p_{bt}}} = \mathrm{w}_{bt}\gamma + \omega_{bt} \tag{42}$$

This is the equation I take to the data to estimate the price collusion parameters.

4.2.2 First Stage Product Decision

In stage 1, firm b chooses subset of J products to offer in market t, taking into account the effect that the choice of products has on the second stage pricing decision. Firm b chooses flavors to solve:

$$\max_{d_{b1},\ldots,d_{bT}} \sum_{t=1}^{T} \mathbb{E}_{\xi}\left[\pi_{bt}\left(d_{bt}, d_{-bt}, p_t(d_{bt}, d_{-bt}; \theta_2)\right)\right] + \theta_{1,b} \sum_{t=1}^{T} \mathbb{E}_{\xi}\left[\pi_{-bt}\left(d_{bt}, d_{-bt}, p_t(d_{bt}, d_{-bt}; \theta_2)\right)\right] \tag{43}$$

Unlike in the theoretical model, the characteristics defining the flavors are not continuous, so the equilibrium choices of flavors (d^*_{bt}, d^*_{-b}) are not defined by a set of first order conditions. However, a necessary condition for a SPNE is that each firm b could not be made better off by unilaterally deviating to an alternative flavor choice d'_{bt}. Thus, the following set of inequalities have to hold at an equilibrium.

$$\begin{aligned} &\sum_{t=1}^{T} \mathbb{E}_{\xi}\left[\pi_{bt}\left(d^*_{bt}, d^*_{-bt}, p_t(d^*_{bt}, d^*_{-bt}; \theta_2)\right)\right] + \theta_{1,b} \sum_{t=1}^{T} \mathbb{E}_{\xi}\left[\pi_{-bt}\left(d^*_{bt}, d^*_{-bt}, p_t(d^*_{bt}, d^*_{-bt}; \theta_2)\right)\right] \geq \\ &\sum_{t=1}^{T} \mathbb{E}_{\xi}\left[\pi_{bt}\left(d'_{bt}, d^*_{-bt}, p_t(d'_{bt}, d^*_{-bt}; \theta_2)\right)\right] + \theta_{1,b} \sum_{t=1}^{T} \mathbb{E}_{\xi}\left[\pi_{-bt}\left(d'_{bt}, d^*_{-bt}, p_t(d'_{bt}, d^*_{-bt}; \theta_2)\right)\right] \; \forall d'_{bt} \neq d^*_{bt}, b, t \end{aligned} \tag{44}$$

where $p_t(d^*_{bt}, d^*_{-bt}; \theta_2)$ are the observed prices and $p_t(d'_{bt}, d^*_{-bt}; \theta_2)$ are the prices chosen in the second stage given the alternative set of flavors but holding the pricing collusion parameters fixed.[27] I take these inequalities to the data in order to estimate the product collusion parameters $\theta_1 = (\theta_{11}, \theta_{12})$ and the retail costs R_{bt}.

5 Data

My data come from the Nielsen Supermarket Scanner Dataset. This dataset contains weekly price and quantity data for every barcode sold in a subset of supermarkets, mass-merchandizers, and convenience stores in the United States from 2006-2013. As will be discussed below, I use weekly-level data to estimate demand. However, like DMS (2009), I use monthly data to estimate the supply model. Thus, I generate a monthly dataset for 5,377 stores. To be included in this sample, the store must be classified as a supermarket by Nielsen. It must also report positive sales of both Ben & Jerry's and Häagen-Dazs in each week of the sample. I also exclude any products sold by either Ben & Jerry's or Häagen-Dazs that are not full-fat ice cream. I further restrict attention to pints, excluding quarts from the analysis.[28]

Monthly quantities are formed by adding the quantity purchased for each flavor offered in the store across weeks. As was mentioned in Section 4.1, I am assuming that the brands charge one

[27] The rationale for why the collusion parameters can be held fixed is discussed in Section 2.2.4.

[28] Very few supermarkets sell quarts of Ben & Jerry's and Häagen-Dazs and the brands only produce them for a small subset of flavors.

price for all flavors in each market. However, in the data I observe different prices for flavors offered by a brand in a given supermarket-week. Nielsen reports the average price for the product sold in a given week. Thus, the reason the average prices differ within a brand appears to be that prices change midweek, causing the average prices for the most purchased products to differ from the average prices of those less purchased. Therefore, I reset the prices to equal the median price reported for a brand in a given supermarket week.[29] I then take a weighted average of the weekly prices to generate the monthly price.

Because aggregating to the month level eliminates much of the variation in the data that will identify the substitution patterns, I create a weekly-level dataset to estimate demand. It is not computationally feasible to include all 5,377 stores in the estimation, so I restrict attention to the 39 stores that sold the most pints of Ben & Jerry's and Häagen-Dazs during my sample. I choose the largest stores for two reasons. First, the largest stores carry the most flavors, allowing me to estimate brand-flavor fixed effects for all the flavors present in the larger sample. Second, a feature of scanner data is that products that do not sell any units during a week are not included in the dataset. Because ice cream sales are likely to be correlated with market size, this problem will be mitigated for the largest stores. I correct for the missing data problem by assuming that a flavor must appear for four consecutive weeks after any appearance in a store. After filling in the missing flavors, I have to adjust their quantities as the logit model does not allow a product to have a market share of zero. To do so, I increase the quantity of all products, including the outside option, by one before computing the market share.[30] Like in the monthly sample, I reset the prices in the data to the median price charged by the brand in the store-week.

I also need to compute the market size, which I estimate from data on weekly milk purchases for the stores in my dataset. I define the market size based on milk sales for a few reasons. Most importantly, because it is both a staple good and perishable, consumers who drink milk are likely to purchase it each week. Also, it is a dairy product, so its demand should be correlated with demand for ice cream. Using the annual per capita consumption of fluid milk reported by the US Department of Agriculture, I am able to estimate the number of customers shopping in a store in a given week from its milk sales.[31] I then average the weekly sales across all weeks in my data and use the average number of consumers as the weekly market size in a store, fixing the market size of a supermarket for the entire sample.[32]

I obtain cost data from a variety of sources. In particular, I collect data on shifters of the marginal

[29] This rules out the possibility that the brands can offer sales for specific flavors.

[30] Gandhi, Lu, and Shi (2013) propose more sophisticated techniques for dealing with this problem.

[31] $\frac{\text{number of customers}}{\text{week}} = \frac{\text{oz. milk sold}}{\text{week}} * \frac{\text{52weeks}}{\text{1year}} * \frac{\text{1lb.}}{\text{128oz.}} * \left(\frac{\text{lbs. milk consumed}}{\text{per person year}}\right)^{-1}$

[32] DMS (2009) define the market size using the total sale of ice cream sold in the same size containers as their inside good. In my context, this would correspond to a market size based on the total sale of pints. I do not use this definition for two reasons. First, it does not capture the consumers who choose to purchase other sizes of ice cream or who choose not to buy ice cream at all. Second, the estimation routine is considerably slower if the market is defined with the sales of pints than if it is defined with milk sales. On average, Ben & Jerry's and Häagen-Dazs account for a large fraction of the pints sold in supermarkets but are purchased by only a small fraction of the total number of customers shopping in a supermarket. Dube, Fox and Su (2012) show that the speed of the convergence of the BLP contraction increases as the share of the outside option increases. I plan on checking the robustness of the results to alternative definitions of the market size in future work.

cost of production, distribution, and retail. Determinants of production costs include average annual wages paid to grocery workers by state, which I obtain from the Bureau of Labor Statistics Quarterly Census of Employment and Wages[33] and the average monthly price of electricity for industrial users for each state.[34] From 2006-2013, Ben & Jerry's manufactured ice cream in three locations: St. Albans VT; Waterbury, VT; and Henderson, NV; while Häagen-Dazs had two manufacturing plants: Laurel, MD and Tulare, CA. I assume that the ice cream sold in each supermarket was produced in the nearest manufacturing plant. Nielsen reports the FIPS state and county code for each store in the dataset. I compute the distances from a store to each plant using the latitude and longitudes for each FIPS county obtained from the 2010 US Census Gazetter.[35] I then assign the production costs for the nearest plant to each store. To proxy for the cost of distribution, I obtain the average regional on-highway diesel prices in dollars per gallon by month from the US Energy Information Administration.[36] For each store, I choose the diesel prices in the region in which the store is located and then interact this with the distance to the nearest manufacturing plant to measure total fuel costs. At the retail level, I get average weekly wages over a given quarter for food manufacturing workers by state from the BLS Quarterly Census of Employment and Wages. I also get the average monthly price paid for electricity by commercial users in each state from the EIA.

6 Identification And Estimation

6.1 Demand Parameters

The main threat to identification of the demand parameters in my model, as is the case generally with demand estimation, is that firms might choose their products and set their prices based on realizations of unobservable determinants of utility. To ameliorate this issue, I follow Nevo (2001) and include product, store, month, and year fixed effects (ξ_{bj}, ξ_s, ξ_m, and ξ_y). These fixed effects control for all unobservables that are constant either within a product, a store, a month, or a year. Thus, the source of any endogeneity is limited to the set of unobservables that vary across these dimensions, denoted as ξ_{bjt}.

Endogeneity resulting from selection on ξ_{bjt} is ruled out with the timing assumption made in Section 4.1. In each market t, firms choose their products and prices before any demand shocks specific to market t are realized. The inclusion of fixed effects requires the distributional assumption in Section 4.1: the ξ_{bjt} are *iid* across all products within a market and across markets. These assumptions taken together yield the following identification assumption:

$$\mathbb{E}[\xi_{bjt}|d_{bt'}, d_{-bt'}, p_{bt'}, p_{-bt'}] = 0 \qquad \forall b, j, t, t' \tag{45}$$

I estimate the parameters in the market share equation using continuously updated GMM fol-

[33]http://data.bls.gov/cgi-bin/dsrv?en
[34]http://www.eia.gov/electricity/data/browser/
[35]https://www.census.gov/geo/maps-data/data/gazetter2010.html
[36]https://www.eia.gov/dnav/pet/pet_pri_gnd_dcus_nus_w.htm

lowing the BLP algorithm. To simulate the integral in the market share equation, I use 1000 Halton draws. In order to identify the standard deviations of the random coefficients, I need a set of instruments. Because these standard deviations serve to distinguish the estimated substitution patterns from those implied by the logit model, the variation in the instruments should induce consumer substitution. Two ways to induce consumer substitution are to change the characteristics of the set of competing products and to change the relative prices charged for those products. I exploit both sources of variation to identify the standard deviation of the random coefficients.

For product j produced by brand b in market t, the instruments I use to identify the standard deviation on the constant and the brand indicator for Ben & Jerry's include the total number of flavors offered in the market, the number of flavors offered by brand b, and the price ratio $\frac{p_{-bt}}{p_{bt}}$.[37] As the number of products offered in the market increases, consumers are induced to substitute from the outside option to the inside goods. Likewise, when one brand offers an additional product or lowers its price relative to its rival, consumers are induced to substitute to that brand. To identify the standard deviation of the coefficient on base flavors and mix-ins, I follow Gandhi and Houde (2015) and proxy for the amount of competition faced by each flavor with respect to that characteristic using two instruments. First, I measure the number of other flavors sold by brand b in the market that share the same value for the characteristic as product j. Then, I measure the number of flavors produced by $-b$ that have the same value of the characteristic as product j. Because the products and prices in market t are chosen before ξ_t is realized, all these variables are valid instruments.

Implementing the BLP algorithm provides estimates for all parameters except $\bar{\beta}$, the means of the random coefficients. Given that the store fixed effects are estimated via the within estimator, it is not possible to identify the mean of the random coefficient on the constant. However, Nevo (2001) shows that the mean values of the random coefficients on the observed product characteristics can be recovered by regressing the estimated product fixed effects on the characteristics defining those products. Specifically let

$$\xi_{bj} = \bar{\beta}\tilde{X}_{bj} + \mu_{bj} \tag{46}$$

where $\tilde{X}_{bj}$ includes the brand dummy for Ben & Jerry's, the indicators for the four base flavors, and the number of mix-ins. I assume $\mathbb{E}[\mu_{bj}|X_{bj}] = 0$ and estimate $\bar{\beta}$ as:

$$\hat{\bar{\beta}} = (\tilde{X}'V_d^{-1}\tilde{X})^{-1}\tilde{X}'V_d^{-1}\hat{\xi}_{bj} \tag{47}$$

where V_d is the covariance matrix of the estimated product fixed effects.

As mentioned in the data section, I estimate the demand parameters using weekly data for the 39 largest supermarkets by sale of Ben & Jerry's and Häagen-Dazs. I estimate the supply side using monthly data for 5,377 stores. I impose that the demand parameters estimated from the subset of stores hold across all stores. This is reasonable because the subset is geographically

[37] I can form the price ratio because there is only one price for each brand in each market.

diverse. Furthermore, I have no reason to think consumer tastes for ice cream vary much across the population. Then, holding these parameters fixed, I can recover the store fixed effects and the demand shocks ξ_{bjt} for the larger sample. This allows me to non-parametrically estimate the empirical distribution of ξ_{bjt} which I use to simulate firm expectations.[38]

6.2 Pricing Stage Parameters

I can either directly observe or simulate all variables in the pricing equation, (42) except the unobserved cost shifter ω_{bt}. I assume ω_{bt} is *iid* both across brands within a market and across markets. Thus to estimate the price collusion parameters, one could imagine rearranging (42) by moving the markups to the right hand side of the equation. At first glance, it might appear that the price collusion parameters could then be directly estimated by non-linear least squares. However, the brand-level market shares and price derivatives in the markups are explicit functions of prices, and are therefore correlated with the unobservable cost shifters ω_{bt}. As such, a set of instruments Z_{bt} for the brand-level market shares and price derivatives is required. Bresnahan (1982) and Berry and Haile (2014) show that firm conduct can be separately identified from marginal cost via the use of instruments that rotate and shift the demand curve. In particular, Berry and Haile (2014) highlight "variation in the number of competing firms, the set of competing goods, characteristics of competing products, or costs of competing firms" (pg. 1779). While the number of competing firms does not change in my model, I exploit the other three sources of variation.

Because I am instrumenting for brand-level market shares and price derivatives, Z_{bt} should contain variables defined over the set of products offered by each brand in market t as opposed to the individual products that comprise those sets. Thus, I include eight market specific brand-level instruments in Z_{bt}: the number of products offered by each brand, the average popularity of the flavors in each brand's product set, three variables that measure the distance in characteristic space between those product sets, and the observed cost shifters of the rival brand. I proxy for the popularity of a flavor in a given market with its brand-flavor fixed effect, ξ_{bj}. I use one variable to measure the distance in characteristic space between the product sets with respect to the base flavors (vanilla, chocolate, coffee, fruit). Specifically, for each base, I compute the absolute difference in the number of flavors offered by each brand containing that base flavor. I then add these distances across bases. Then, to capture the distance with respect to the number of mix-ins I include two instruments, one for each brand, which measure the average number of mix-ins per flavor offered by that brand in each market.

In theory, these instruments should be relevant because variation in each should induce consumer substitution across brands, affecting both the brand-level market shares and price derivatives. By offering an additional product or increasing the popularity of the set of products it currently offers, a brand can induce substitution from the rival brand and the outside option to its own flavors.

[38] In my estimates, I ignore issues of heteroskedasticity in ξ_{bjt}. However, it is likely that the variance depends on, amongst other variables, the size of the supermarket. I plan on addressing this in future work by allowing the variance of ξ_{bjt} to depend on the market size.

Likewise, an increase in the distance between the brand's offerings or a rival's marginal cost should affect the prices charged by the brands, impacting both the brand shares and price derivatives. The instruments are also exogenous, which is guaranteed by the timing and distributional assumptions in the model. In particular, I assume that the set of products to be offered in market t is fixed at the time firm b makes its pricing decision. Furthermore, the brand-flavor fixed effects are assumed known to each firm, which chooses its price conditional on their value. Likewise, the observed cost shifters for firm b in market t are assumed uncorrelated with the unobserved determinants of c_{-bt}.

I make one additional identifying assumption, that the observed cost shifters for firm b in market t are uncorrelated with the unobserved determinants of c_{bt}. Thus, for each brand in each market the following moment condition holds:

$$\mathbb{E}[\omega_{bt}(\theta_2, \gamma) | Z_{bt}, \mathrm{w}_{bt}] = 0 \qquad \forall b, t \tag{48}$$

where $\theta_2 = (\theta_{21}, \theta_{22})$. With this moment, both the price collusion parameters and the parameters on the cost shifters could be directly estimated via GMM. Instead, I replace Z_{bt} with the optimal instruments Z^*_{bt} to improve efficiency, where Z^*_{bt} defined as:

$$Z^*(Z_{bt}, \mathrm{w}_{bt}) = \mathbb{E}\left[\frac{\partial \omega_{bt}}{\partial \rho} \middle| Z_{bt}, \mathrm{w}_{bt} \right] \tag{49}$$

where $\rho = (\theta_2, \gamma)$. Thus, I consider the following sample moment:

$$G(\theta_2, \gamma) = \frac{1}{2T} \sum_t \sum_b \omega_{bt}(\theta_2, \gamma) \hat{Z}^*_{bt}(\theta_2, \gamma, Z_{bt}, \mathrm{w}_{bt}) \tag{50}$$

where $\hat{Z}^*$ is a continuously updated estimate of Z^*. With this sample moment, I estimate the parameters in (42) via GMM such that:

$$(\hat{\theta}_2, \hat{\gamma}) = \arg\min_{\theta_2, \gamma} \; G' W G \tag{51}$$

where W is the inverse of the finite sample variance of $G(\theta_2, \gamma)$.

To simplify estimation, I solve the maximization problem sequentially. For a given guess of θ_2, I compute the left hand side of (42). Note that this requires me to first simulate the expectations in the markup using random draws from the empirical distribution of ξ_{bjt}.[39] Then I regress the left hand side on the cost shifters w_{bt} which yields:

$$\hat{\gamma}(\theta_2) = \left(\mathrm{w}'\mathrm{w} \right)^{-1} \left(\mathrm{w}'\omega(\theta_2) \right) \tag{52}$$

[39]In my estimates, I currently set the ξ_{bjt} to their mean value, 0, when simulating the expectation as opposed to taking S draws from the distribution. This is done for computational reasons. I have compared these results to estimates obtained from a small number of draws and find no meaningful difference. However, I plan on relaxing this in future work.

My estimator for θ_2 is

$$\hat{\theta}_2 = \arg\max_{\theta_2} \; G\Big(\theta_2, \hat{\gamma}(\theta_2)\Big)' W G\Big(\theta_2, \hat{\gamma}(\theta_2)\Big) \tag{53}$$

Given estimates of the price collusion parameters, I can now estimate the product collusion parameters and the market level fixed costs.

6.3 Product Stage Parameters

Let the equilibrium product choice be represented as $d_t^* = (d_{bt}^*, d_{-bt}^*)$. Also, let $d_t' = (d_{bt}', d_{-bt}^*)$ indicate an alternative product choice for firm b in market t while keeping its rival's product choice fixed at d_{-bt}^*. For any function of the product choices, $f(d_{bt}, d_{-bt})$, define:

$$\Delta f(d^*, d') = f(d^*, p(d^*; \theta_2)) - f(d', p(d'; \theta_2)) \tag{54}$$

With this notation, the SPNE conditions defined by (44) can be rewritten as

$$\sum_{t=1}^{T} \mathbb{E}_\xi \Big[\Delta \pi_{bt}(d_t^*, d_t', \cdot)\Big] + \theta_{1,b} \sum_{t=1}^{T} \mathbb{E}_\xi \Big[\Delta \pi_{-bt}(d_t^*, d_t', \cdot)\Big] \geq 0 \qquad \forall d_{bt}' \neq d_{bt}^*, b, t \tag{55}$$

I do not observe the firms' expectations, so it is not possible to evaluate the above inequality directly. However, because I have estimates of the demand parameters, marginal costs, and price collusion parameters, I can simulate the expected profits for both firms up to the fixed cost parameter R_{bt} for any possible choice of products d_t.

$$r_{bt}(d_t) = \frac{1}{S} \sum_s M_t(p_t(d_t, \theta_2) - c_{bt}) \sum_{j=1}^{J} d_{bjt} s_{bjt}(d_t, p_t(d_t, \theta_2), \xi_{st}) \tag{56}$$

To compute $r_{bt}(d_t)$, I draw S vectors from the empirical distribution of ξ_{bjt}, where the length of each draw ξ_{st} corresponds to the number of products in market t. Then, for each draw ξ_{st}, I find the optimal set of prices for both firms to charge ($p_t(d_t, \theta_2)$) by solving for a fixed point to the pricing first order conditions for both firms in market t.[40] With these prices, I can then evaluate the market shares and the profits given ξ_{st}. Averaging the profits across the draws gives an approximation to firm b's expected variable profits.[41]

[40] As was mentioned above, the price collusion parameter can be held fixed given an alternative choice of products. See Section 2.2.4 for further discussion

[41] In the estimates I present, I simplify computation in two ways. First, instead of drawing from the distribution of ξ_{bjt} to simulate the firm's expectations, I set all ξ_{bjt} to their mean value 0. Secondly, when computing the expected profits for an alternative set of products, I use the observed price instead of solving for the firm's optimal price. I plan on removing these restrictions in future work.

The expected profits earned by firm b in market t are given as:

$$\mathbb{E}[\pi_{bt}] = M_t(p_{bt} - c_{bt})\sum_{j=1}^{J} d_{bjt}\mathbb{E}_\xi[s_{bjt}] - R_{bt} \tag{57}$$

Following Pakes (2010), (57) can be rewritten as

$$\mathbb{E}[\pi_{bt}] = r_{bt} + v_{1bt} - v_{2bt} \tag{58}$$

where v_{1bt} is mean 0 error term which includes both simulation and measurement error

$$v_{1bt}(d_t) = M_t(p_{bt} - c_{bt})\sum_{j=1}^{J} d_{bjt}\mathbb{E}_\xi[s_{bjt}] - r_{bt} \tag{59}$$

and v_{2bt} is structural error

$$v_{2bt}(d_t) = R_{bt} \tag{60}$$

With this notation, the SPNE necessary condition can be rewritten as a function of r_{bt}

$$\begin{aligned}\sum_{t=1}^{T}\Delta r_{bt}(d_t^*, d_t') + \theta_{1,b}\sum_{t=1}^{T}\Delta r_{-bt}(d_t^*, d_t') + \sum_{t=1}^{T}\Delta v_{1bt}(d_t^*, d_t') + \theta_{1,b}\sum_{t=1}^{T}\Delta v_{1-bt}(d_t^*, d_t') \\ - \sum_{t=1}^{T}\Delta v_{2bt}(d_t^*, d_t') - \theta_{1,b}\sum_{t=1}^{T}\Delta v_{2-bt}(d_t^*, d_t') \geq 0 \qquad \forall d' \neq d^*\end{aligned} \tag{61}$$

The two parameters to be estimated are the product collusion parameters θ_1 and the market fixed cost R_{bt}. In order to separately identify the two parameters, I rely on an assumption made in Section 4.2: R_{bt} depends only on the number of flavors sold in market t, not the identity of those flavors or the quantities sold. With this assumption, I consider alternative product sets for firm b which differ from the observed choices by one product. There are three ways to construct these alternative product choices: firm b could add a product to d_t^*, remove a product from d_t^*, or replace one product in d_t^* with a product not in d_t^*. I will show that θ_1 is set identified and can be estimated by constructing inequalities with "replacement" moments. Once I estimate θ_1, the "add" and "remove" moments can be used to identify and estimate R.

I first consider identification and estimation of θ_1. In every possible "replacement" moment, the number of products offered in the market is unchanged. Because R only depends on the number of products offered, it is the same for the observed product choice d^* and the alternative product choice d'. Because of this, $\Delta v_{2bt}(d_t^*, d_t') = 0$ for each firm b in each market t and the structural error differences out of the inequality. If I divide both sides of the inequality by the number of markets

T, the inequality becomes:

$$\overline{\Delta r}_b(d^*, d') + \theta_{1,b}\overline{\Delta r}_{-b}(d^*, d') + \underbrace{\overline{v}_{1b}}_{\to_p 0} + \theta_{1,b}\underbrace{\overline{v}_{1-b}}_{\to_p 0} \geq 0 \tag{62}$$

where the bar above a variable represents the average across markets. Notice that since v_{1bt} is *iid* mean 0, as $T \to \infty$, $\overline{v}_{1b}$ and $\overline{v}_{1-b} \to_p 0$. Asymptotically, the inequality can be written as:

$$\overline{\Delta r}_b(d^*, d') + \theta_{1,b}\overline{\Delta r}_{-b}(d^*, d') \geq 0 \tag{63}$$

Rearranging this inequality permits me to construct bounds on θ_{1b}.

$$\theta_{1,b} \lessgtr -\frac{\overline{\Delta r}_b}{\overline{\Delta r}_{-b}} \tag{64}$$

There are four cases summarized in the following table.

Table 1
Possible Bounds on Product Parameters

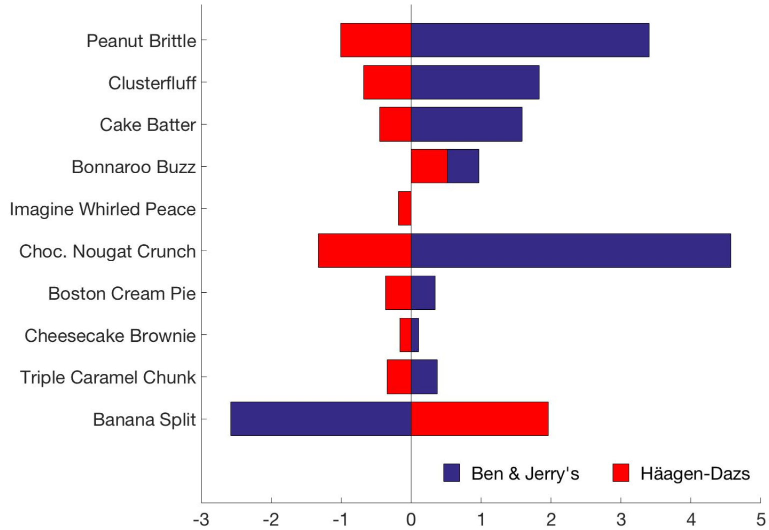
Case	$sgn(\overline{\Delta r}_b)$	$sgn(\overline{\Delta r}_{-b})$	Bound	sgn(Bound)
1	< 0	> 0	Lower	> 0
2	> 0	< 0	Upper	> 0
3	> 0	> 0	Lower	< 0
4	< 0	< 0	Upper	< 0

In the first case, the switch results in higher expected profits for brand b ($\overline{\Delta r}_b > 0$) and lower expected profits for its rival ($\overline{\Delta r}_{-b} > 0$). In the absence of collusion, firm b would offer d'_b instead of d^*_b. Thus, firm b must be internalizing the impact of its product choice on its rival by at least the amount needed to prevent it from offering d'_b. Thus, this case provides a positive lower bound on the value of $\theta_{1,b}$. In the second case, the switch results in lower expected profits for brand b ($\overline{\Delta r}_b < 0$) and higher expected profits for its rival ($\overline{\Delta r}_{-b} < 0$). This case shows the limit on the degree to which firm b internalizes the impact it has on its rival, providing an upper bound on the collusion parameter. In the third case, both firms are hurt by the switch ($\overline{\Delta r}_b, \overline{\Delta r}_{-b} < 0$). This switch is uninformative about the value of the collusion parameter, providing a negative lower bound. Finally, in the fourth case, both firms would be made better off as a result of the switch, providing a negative upper bound. This case would make sense during a punishment period in which firms forego mutually beneficial behavior. I assume the firms are choosing actions on the equilibrium path, so the first two cases provide the bounds on the collusion parameters.

Because I include brand-flavor fixed effects in the demand system, I can only consider alternative product choices which switch a flavor produced by brand b in 2013 with one it produced earlier in my sample. Though this limits the inequalities I can construct, this restriction allows me to control for a competing rationale for the observed product choices, namely, that consumers prefer the chunky flavors produced by Ben & Jerry's and the smooth flavors produced by Häagen-Dazs. During 2013, Ben & Jerry's offered 39 of its non-seasonal flavors for the entire year. Häagen-Dazs offered 32 such flavors. Meanwhile, from 2006-2012, Ben & Jerry's produced 19 non-seasonal flavors that were not sold in 2013 while Häagen Dazs produced 26 such flavors.[42] Therefore, I am able to construct 741 alternative product choices for Ben & Jerry's and 832 alternative product choices for Häagen-Dazs.

FIGURE 5

AVERAGE CHANGE IN PROFITS FOR EACH BRAND BY REPLACING TEN WORST SELLING BEN & JERRY'S FLAVORS WITH CHOCOLATE



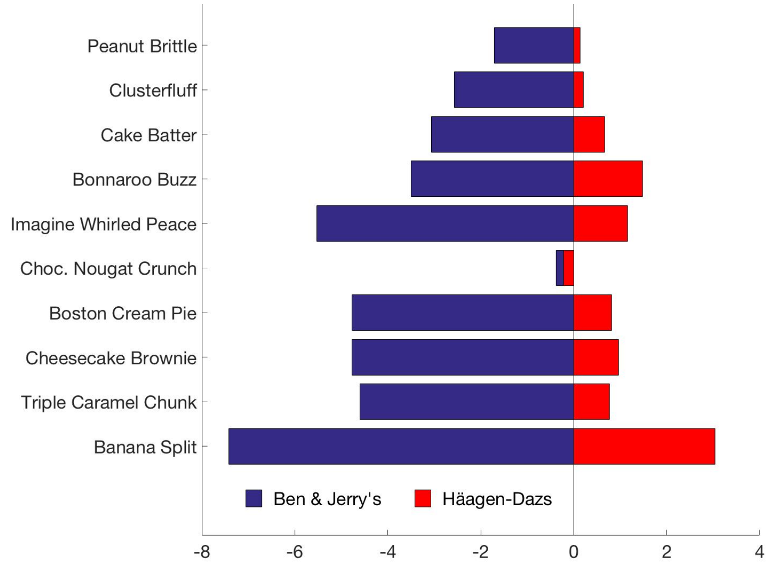
From the entire set of alternative product choices for each brand b, a subset will be informative about the lower bound of $\theta_{1,b}$. This subset contains all switches that increase the profits earned by brand b while lowering the profits earned by its rival. One such switch could be generated by replacing an unpopular flavor produced by brand b with its own version of a popular flavor produced by $-b$. I illustrate this idea by performing switches in which each of the ten worst selling Ben & Jerry's flavors during 2013 are replaced with chocolate. Figure 5 shows the average change in expected profits for both Ben & Jerry's and Häagen-Dazs resulting from each pairwise switch. These switches result in an increase in expected profits for Ben & Jerry's and a decrease in expected profits for Häagen-Dazs in seven of the ten cases, and thus would be informative about the lower bound of $\theta_{1,b}$.

Meanwhile, the subset of alternative product choices that hurt brand b but help its rival will be informative about the upper bound of $\theta_{1,b}$. This can be achieved by replacing a popular flavor produced by brand b with one of its previously sold unpopular flavors. In Figure 6, I consider the average effect on expected profits by replacing Ben & Jerry's ten worst selling flavors with Black

[42] I avoid moments constructed from seasonal flavors for three reasons. First, I do not estimate brand-flavor fixed effects separately for each season. Thus, the brand-flavor fixed effect I estimate is only appropriate in the season the flavor is offered. Furthermore, because Ben & Jerry's and Häagen-Dazs know these flavors have a limited releases, they may offer these flavors with dynamic considerations, such as experimentation. Also, in my dataset, seasonal flavors are sold with the same barcode. Therefore, I have to arbitrarily declare the date on which the identity of the seasonal flavors changes.

and Tan.[43] The effect of each pairwise switch is to lower Ben & Jerry's profits and increases the profits of Häagen Dazs in all but one case, making them informative about the upper bound of $\theta_{1,b}$.

FIGURE 6
AVERAGE CHANGE IN PROFITS FOR EACH BRAND BY REPLACING
TEN WORST SELLING BEN & JERRY'S FLAVORS WITH BLACK & TAN



To estimate the bounds on $\theta_{1,b}$, I generate the following sample moment for each possible alternative product choice $k = 1, \ldots, K_b$:

$$\bar{m}_{bk}(\theta_{1,b}) = \overline{\Delta r}_b(d^*, d'_k) + \theta_{1,b}\overline{\Delta r}_{-b}(d^*, d'_k) \tag{65}$$

I then aggregate across the moments. Following Pakes et al. (2011), the estimator for set of collusion parameters $\Theta_{1,b}$ is given as:

$$\Theta_{1,b} = \arg\min_{\theta \in \Theta} \sum_{k=1}^{K_b} \left[\frac{\bar{m}_{bk}(\theta_{1,b})}{\sigma_{bk}(\theta_{1,b})} \right]_-^2 \tag{66}$$

where $[x]_- = \min\{0, x\}$. Following Andrews and Soares (2010), each moment is weighted by σ^2_{bk}, an estimate of the asymptotic variance of $n^{1/2}\bar{m}_{bk}(\theta)$ where:

$$\sigma^2_{bk}(\theta) = T^{-1} \sum_{i=1}^{T} (m_{bkt}(\theta) - \bar{m}_{bk}(\theta))^2 \tag{67}$$

With this estimator, I estimate the bounds on the parameter set as:

$$\hat{\underline{\theta}}_{1,b} = \min(\Theta_{1,b}) \qquad \hat{\bar{\theta}}_{1,b} = \max(\Theta_{1,b}) \tag{68}$$

[43]This flavor was created by Ben & Jerry's to taste like the Black and Tan drink, which is an ale mixed with a stout. It was offered in 2006 and 2007 and was unpopular. It can now be found in the Flavor Graveyard.

In practice, I find $\underline{\hat{\theta}}_{1,b} = \hat{\bar{\theta}}_{1,b}$ for both brands.

Given estimates of the collusion parameters, it is possible to find the market-level fixed cost. Under the assumption that $R_{bt} = \eta_b N_{bt}$, I can generate two types of alternative product sets, those in which firm b adds a product to d^*_{bt} and those in which firm b removes a product from d^*_{bt}. Alternative sets created by adding an additional product result in the following inequalities:

$$\bar{r}_b(d^*, d'_k) + \hat{\theta}_{1,b}\bar{r}_{-b}(d^*, d'_k) + \eta_b \geq 0 \tag{69}$$

Adding a flavor provides a lower bound on the per flavor fixed cost η_b. Since firm b chose not to include this flavor in equilibrium, the fixed costs have to be large enough to justify this decision. Meanwhile, the following inequalities are derived by removing a product from d^*_{bt}:

$$\bar{r}_b(d^*, d'_k) + \hat{\theta}_{1,b}\bar{r}_{-b}(d^*, d'_k) - \eta_b \geq 0 \tag{70}$$

These moments provide an upper bound on η_b. Because firm b chose to offer the removed flavor in equilibrium, the fixed cost cannot exceed the benefit firm b received by offering it.

Given the sets of flavors produced in my dataset, I am able to construct 39 alternative product sets for Ben & Jerry's by removing one flavor and 19 alternative product sets by adding an additional flavor. For Häagen-Dazs, I construct 32 alternative product sets by removing a flavor and 26 alternative product sets by adding a flavor. With the moments that are generated from these alternative product sets, I estimate bounds on η_b and η_{-b} following the procedure described for θ_1.

7 Results

7.1 Demand Parameter Estimates

Table 2 presents the estimates of the demand parameters. Reassuringly, I find the price coefficient α is negative. I also find vanilla has the highest mean utility, followed by chocolate. Furthermore, I find that the standard deviations for the random coefficients on fruit and coffee are considerably higher than those for chocolate and vanilla. This matches my expectation that vanilla and chocolate have broad appeal throughout the population while coffee and fruit tend to be more polarizing. The constant is negative, which results mechanically from the fact that the share of the outside option is large in most markets (∼90%).

Table 2

Estimates Of Demand Parameters

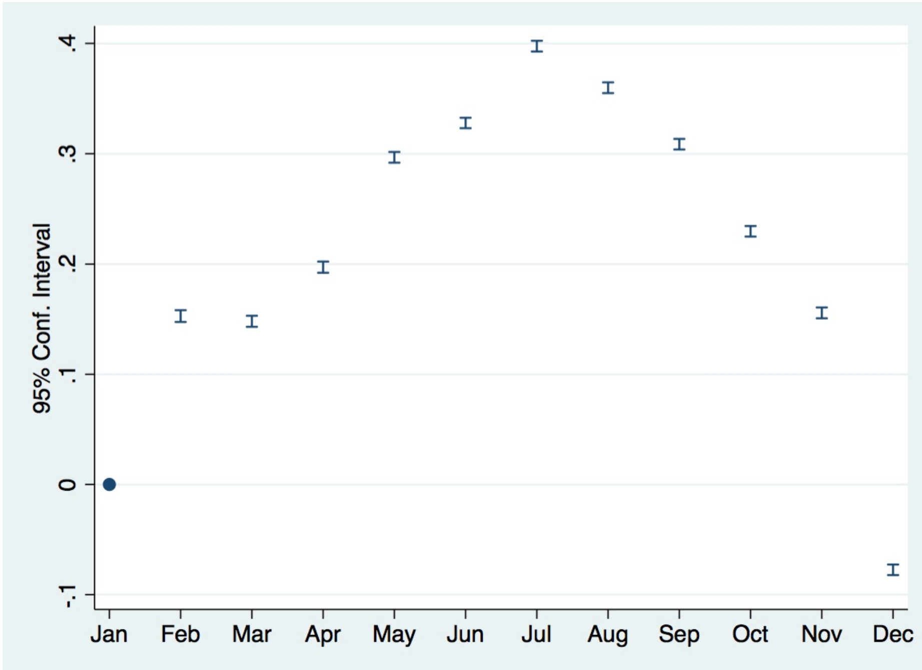
Parameter	Mean	Standard Deviation
$\beta_{constant}$	-11.8350[a]	2.7354***
		(0.0058)
$\beta_{Ben\ \&\ Jerry's}$	-4.9678***	0[b]
	(0.0204)	(0.0193)
$\beta_{vanilla}$	1.2686***	1.8406***
	(0.0111)	(0.0015)
$\beta_{chocolate}$	0.4240***	1.1590***
	(0.0167)	(0.0021)
β_{coffee}	-6.9614***	2.7405***
	(0.0233)	(0.0016)
β_{fruit}	-4.0359***	3.3601***
	(0.0197)	(0.0012)
$\beta_{mix\text{-}ins}$	-0.3254***	0.3246***
	(0.0055)	(0.0004)
α	-0.6340***	—
	(0.0013)	

a. mean value of the random coefficient on the constant is not separately identifiable from the store fixed effects. I have reported the mean of the store fixed effects weighted by the number of observations in each store.

b. σ_{brand} is estimated at boundary. Therefore, standard error not credible.

I also check that the estimated month fixed effects match the seasonal pattern ice cream demand is known to follow. Figure 7 plots the 95% confidence interval for the estimated month fixed effects (ξ_m) where the fixed effect for January has been normalized to zero. The pattern is exactly as expected, consumer preferences for super-premium ice cream increase from February through July and then decrease from July through December.

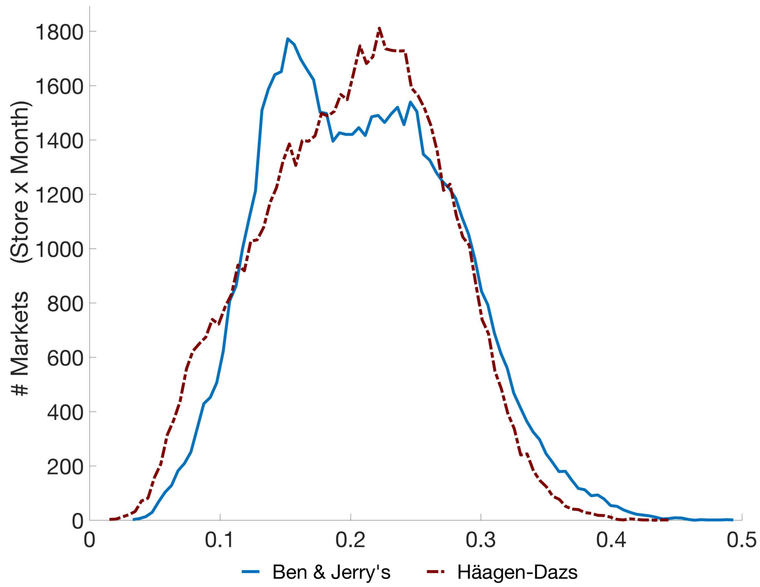
FIGURE 7
ESTIMATED SEASONALITY OF ICE CREAM DEMAND



Plot displays the estimated 95% confidence intervals for month fixed effects (ξ_m). Standard errors for each ξ_m are computed using the standard formula for GMM standard errors with optimal weight matrix. The fixed effect for January has been normalized to 0.

Finally, the standard deviation of the constant plays a particularly important role in determining the substitution patterns. In a standard logit specification, an increase in price causes consumers to substitute to goods proportionally based on their relative market shares. Because the share of the outside option is so large, the logit model would predict that an increase in Ben & Jerry's price would result mostly in substitution to the outside option as opposed to Häagen-Dazs. The fact that the constant has a large standard deviation helps generate more realistic substitution patterns in this model. In Figure 8, I plot a histogram of $\Delta_{b,-b,t}$ for each brand. In an average market, around 25% of the customers who switch as a result of a price increase switch to the rival brand as opposed to the outside option. Though this might seem small, the outside option includes all other brands of ice cream as well as not purchasing any ice cream.

Figure 8
Estimated Cross Brand Substitution



7.2 Pricing Stage Parameters

Parameter estimates from the pricing stage are included in Table 3. I estimate three different specifications. In the first two specifications, I constrain Ben & Jerry's and Häagen-Dazs to have the same price collusion parameter. In the third specification, I allow the price collusion parameters for the two brands to differ. This is my preferred specification and I use these estimates in all subsequent analysis. In all three specifications, I control for ingredient prices using month fixed effects. This assumes that both brands face the same cost of ingredients. In addition, I include cost shifters for the manufacturing, distribution, and retail components of marginal cost. The last two specifications have store fixed effects to control for additional store specific components of marginal cost.

I find strong evidence that Ben & Jerry's and Häagen-Dazs were colluding on price during 2013. Estimates from the first specification suggest that the brands internalize 37% of the externality each imposes on its rival. When I include store fixed effects, this rises to 66%. Allowing the collusion parameters to differ across brands does not substantially alter the estimates, which I would expect given that these are similarly sized firms. Importantly, all three specifications reject zero as the value of the collusion parameters. Though the collusion parameters are informative as to the presence of price collusion, they cannot tell us how far firms are pricing above the Nash prices and, therefore, do not speak directly to the impact on social surplus. Therefore, I perform counterfactual analysis in the next section in which I set the price collusion parameters to 0.

Table 3

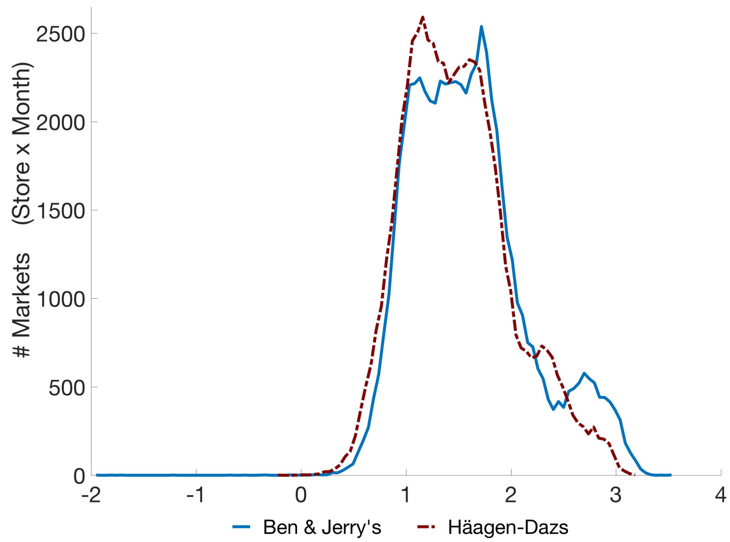
Estimates Of Pricing Stage Parameters

	(1)	(2)	(3)
$\theta_{2,\mathit{Ben\ \&\ Jerry's}}$	0.3743***	0.6590***	0.5366***
	(0.0075)	(0.0034)	(0.0032)
$\theta_{2,\mathit{Häagen\text{-}Dazs}}$	0.3743***	0.6590***	0.8135***
	(0.0075)	(0.0034)	(0.0034)
Cost Shifters:			
Constant	6.2426	3.8228[a]	3.6772[a]
	(0.0012)		
Ben & Jerry's	0.0696	0.0836	0.2111
	(0.0018)	(0.0001)	(0.0018)
Food Manu. Wage	0.1231	0.1060	0.10014
($ '00s/week, quarter x state)	(0.0002)	(0.0002)	(0.0016)
Indus. Elec. Price	-0.0130	$5.82\text{x}10^{-5}$	-0.0035
(cent/kWh, month x state)	(0.0001)	($4.03\text{x}10^{-5}$)	(0.0004)
Diesel	-1.3316	-0.8372	-0.7926
($/gal, month x state)	(0.0003)	(0.0035)	(0.0108)
Diesel x Distance (000's)	0.0326	-0.0411	-0.0416
	(0.0005)	($4.78\text{x}10^{-5}$)	(0.0018)
Commer. Elec. Price	0.0274	0.0506	0.0495
(cent/kWh, month x state)	(0.0001)	(0.0024)	(0.0011)
Grocery Wage	-0.0158	–	–
($ '000s/year, year x county)	($5.24\text{x}10^{-5}$)		
Customers ('000s)	0.0229	–	–
	(0.0002)		
Month FEs	Yes	Yes	Yes
Store FEs	No	Yes	Yes

The marginal cost estimates are both functions of the demand parameters and price collusion parameters. Figure 9 presents the distribution of estimated marginal costs across markets for each brand.

Figure 9

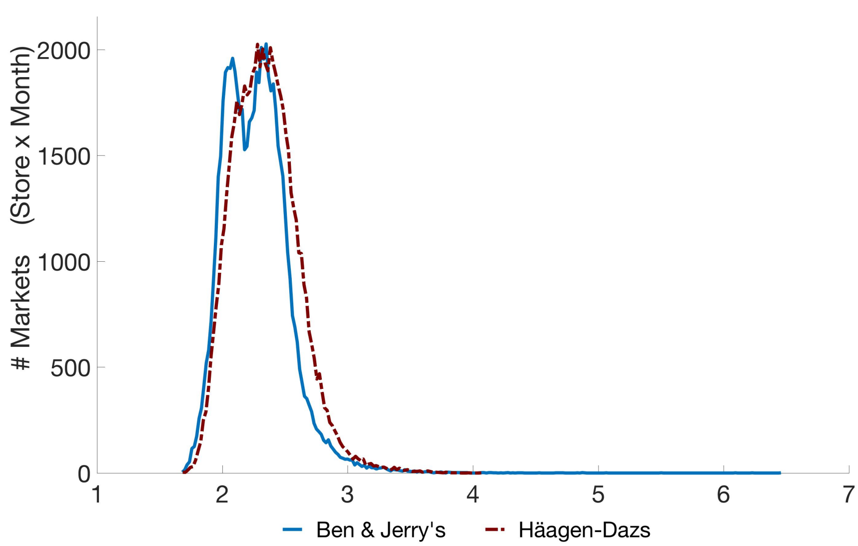
Estimated Distribution Of Marginal Costs (in \$)



An important feature of these distributions is that there is almost no mass below zero. Furthermore, the implied price cost markups seem reasonable. For the vast majority of markets, the markup on a pint of Ben & Jerry's and Häagen-Dazs ranges from \$2 - \$3. Taken together, this evidence helps validate the demand and price parameter estimates.

Figure 10

Estimated Distribution Of Price Cost Markup ($p_{bt} - c_{bt}$)



7.3 Product Stage Parameters

Table 4 presents estimates of the product space collusion parameters θ_1 and the per flavor retail fixed costs η. I find substantial evidence of product space collusion for both brands. Whereas Häagen-Dazs was more accommodating in the pricing stage, I find that Ben & Jerry's internalizes more of its affect on Häagen-Dazs in the product stage than vice versa.

Table 4
Estimates Of Product Stage Parameters

Product Collusion Parameters:	
$\theta_{1,\textit{Ben \& Jerry's}}$	1.7900
$\theta_{1,\textit{Häagen-Dazs}}$	1.5428
Per-Flavor Retail Fixed Cost:	
$\eta_{\textit{Ben \& Jerry's}}$	\$8.39
$\eta_{\textit{Häagen-Dazs}}$	\$10.34

Having obtained estimates of the fixed costs, it is possible to compute the profits the firms expected to earn in each market. Table 5 reports the total expected profits across all markets in my sample as well as the components of those profits. For publicly traded companies, it is easy to validate the estimates in Table 5 as the companies have to list their annual revenue, cost and profit in their 10-K filing to the SEC. Unfortunately, Häagen-Dazs has been a subsidiary of various corporations since 1983 while Ben & Jerry's was acquired by Unilever in 2000, thus there is no recent data to compare.[44]

Table 5
Total Firm Profits And Costs In Sample (in millions)

	Ben & Jerry's	Häagen-Dazs
Total Expected Revenue	\$76.844	\$78.470
Total Expected Variable Costs	\$28.279	\$27.301
Total Fixed Costs	\$13.767	\$13.277
Total Expected Profit	\$34.798	\$37.892

Table 5 provides a compelling interpretation of the estimated fixed costs. The expected variable profits (revenue - variable costs) measure the producer surplus to be split between each brand and the supermarkets. I find that fixed cost payments account for 28.35% of the expected variable profits

[44]Ben & Jerry's was publicly traded before 2000, and it filed a 10-K in 1999. In that file, Ben & Jerry's reported \$213.8 million in profits (measured in 2013 dollars) for 1998. The expected profits in Table 5 are computed from the sale of pints in the 5,377 stores in my sample. According to Bureau of Labor Statistics there were 89,435 supermarkets in US in 2013. Furthermore, in its 10-K, Ben & Jerry's reported that pints accounted for 81% of sales. Extrapolating the expected profits in Table 5 suggests that Ben & Jerry's expected to earn \$714,550,376.82 in profits during 2013. Also in its 10-K, Ben & Jerry's reported that costs accounted for 65.1% of its revenue while profit accounted for the other 34.9% in 1998. I find that costs are 54.72% of its revenue and profits are 45.3% of its revenue in 2013. It is comforting that the estimates are somewhat similar, and that the estimated profits in 2013 are higher than in 1998. In particular, the merger with Unilever likely would have resulted in cost reductions as Ben & Jerry's could utilize Unilever's distribution network.

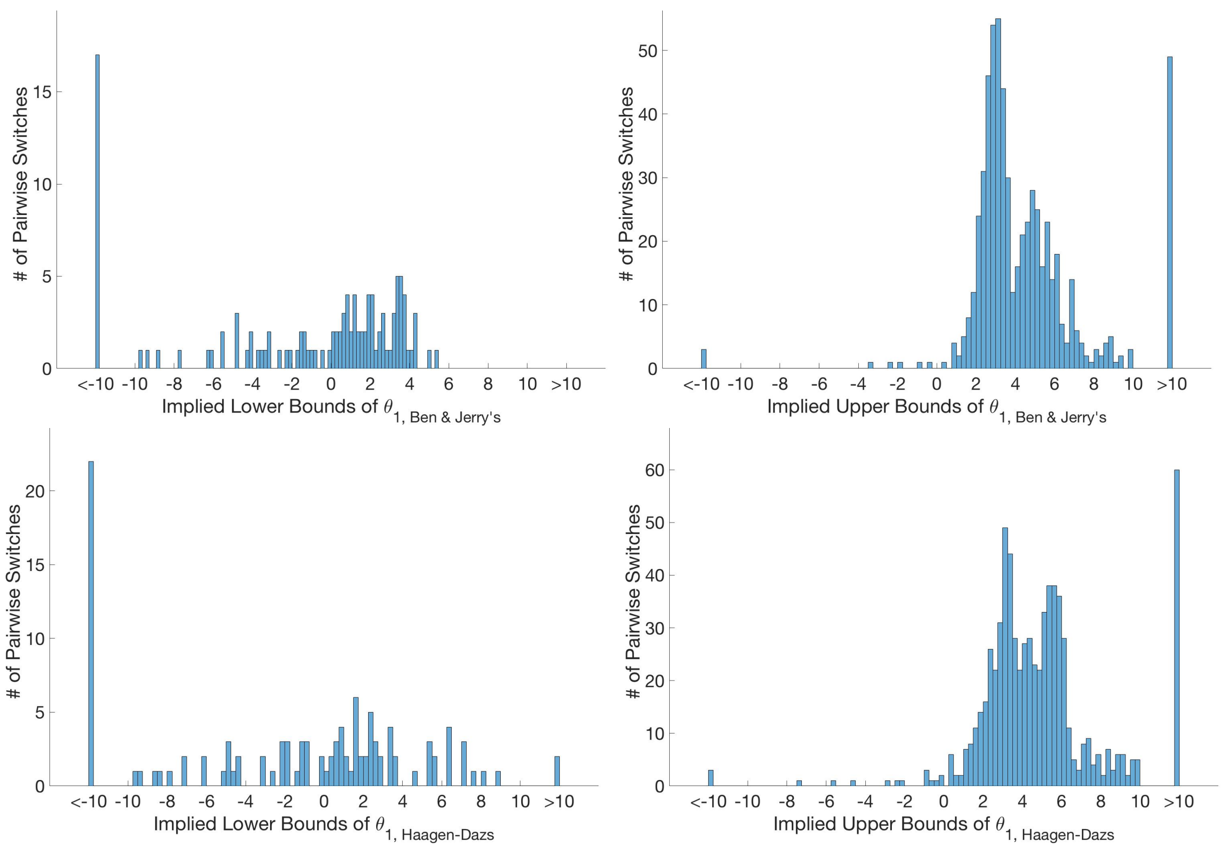
earned by Ben & Jerry's and 25.95% of the expected variable profits earned by Häagen-Dazs. While I have not modeled the bargaining process between the supermarkets and the brands, it seems reasonable that the supermarkets receive between 1/4 and 1/3 of producer surplus.

When considering the product collusion parameters, the reader might be concerned that the estimates lie above one. In the context of my model, these results suggest that the brands place more weight on their rival's profits than their own. To understand why this mechanically occurs, consider the expression in (64). For a given replacement, an implied lower bound of $\theta_{1,b}$ will exceed one if the benefit of the replacement for firm b exceeds the loss to its rival.

Below, I plot the histogram of the bounds implied by the 741 replacement moments generated for Ben & Jerry's and the 832 replacement moments generated for Häagen-Dazs.

Figure 11

Distribution Of Implied Bounds For θ_1



I have separated the histograms into those that imply upper bounds and those that imply lower bounds. It is reassuring that there are not many negative upper bounds. I only estimate negative upper bounds for 1.08% of the 741 Ben & Jerry's deviation moments and 1.94% of the 832 Häagen-Dazs moments. However, it is also apparent that there is still a lot of variability in the moments implied by the bounds. This could partly be the result of misspecification of the profit function or remaining structural error that has not been accounted for. There is another possibility: firms might be highly accommodating in the first stage to signal their commitment to the collusive agreement.

Given that the firms appear to be concluding a product space war from 2006-2012, this might not be unrealistic.

The most likely explanation though is that the demand system is misspecified. In particular, when Ben & Jerry's replaces a flavor with butter pecan, the substitution patterns are not rich enough to fully capture the impact that the replacement would have on Häagen-Dazs sales of butter pecan. Crucially, this misspecification should bias the estimated product collusion parameters upward. Also, the model is missing brand-region fixed effects, which are important given that Ben & Jerry's publicly advocates for liberal political causes. I plan on trying to improve the flexibility of the demand system in future work. Importantly though, the large estimates of θ_1 are suggestive of substantial product collusion between the firms.

8 Counterfactuals

8.1 Effect Of Eliminating Price Collusion

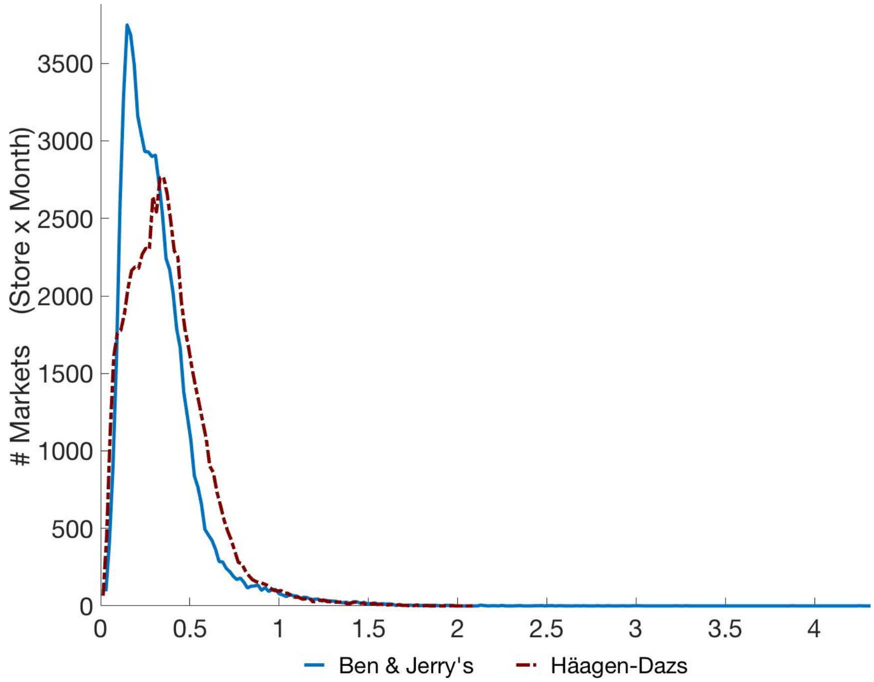
I first want to measure the effect of price collusion on outcomes and welfare ignoring collusion on the product space. These measures will serve as baseline estimates reflecting the current state of the literature. To do so, I estimate the prices firms would have charged in the absence of price collusion, holding the observed set of products fixed. I obtain these price estimates by setting θ_2 to 0, and finding the fixed point of the pricing first order conditions in each market. I simulate the expectations in the first order condition by taking draws from the empirical distribution of ξ_{bjt}.[45]

In Figure 12, I present a histogram of the price increases that result in each market when only price collusion is considered. I find that by colluding on price, Ben & Jerry's was able to charge on average \$0.32 or 9% more for each pint in 2013. The effect on Häagen-Dazs prices was almost identical; on average it charged \$0.38 or 11% more for each pint.

[45] I set all ξ's to 0, their mean value, in order to simplify computation.

Figure 12

Price Increase Above Nash Pricing At Observed Product Choice



The measured effect of price collusion on producer surplus across the markets in my sample is recorded in Table 6. Since the set of products offered is not allowed to adjust, the fixed cost payments to supermarkets do not change, and the effect on producer surplus is fully captured by the change in Ben & Jerry's and Häagen-Dazs profits. Overall producer surplus increases by 4.11% in my sample. Of that increase, 66.32% goes to Ben & Jerry's, 33.68% to Häagen-Dazs. There are 5,377 stores in my sample. According to the Bureau of Labor Statistics there were 89,435 supermarkets in the US in 2013. Extrapolating across all the supermarkets in the US suggests that price collusion yielded a $47,775,994.04 increase in producer surplus in 2013.

Table 6

Total Changes In Producer Surplus From Collusion (in millions)

	Total	Ben & Jerry's	Häagen-Dazs
Total Profits: Price Collusion	$72.689	$34.798	$37.892
Total Profits: No Price Collusion	$69.817	$32.893	$36.924
Benefit of Price Collusion	$2.872	$1.905	$0.967

Following Small and Rosen (1981) and Fan (2013), I estimate the effect of price collusion on consumer welfare via compensating variation. In particular, the expected effect of collusion on consumer surplus can be expressed as follows:

$$\Delta CS = \int \int \frac{V_{it}^{N} - V_{it}^{C}}{\alpha} f(\beta) f(\xi) \mathrm{d}\beta \mathrm{d}\xi \tag{71}$$

where

$$V_{it}^{N} = \ln\left(\sum_{b}\sum_{j=1}^{J} d_{bjt} e^{u_{ibjt}^{N}} + 1\right) \tag{72}$$

and $u_{ibjt}^{C} = U_{ibjt}^{C} - \epsilon_{ijbt}$ is the utility given in (31) minus the idiosyncratic shock evaluated at the observed set of products and prices and where u_{ibjt}^{N} is the same utility measure evaluated at the observed set of products and the prices that would be charged in the absence of price collusion. Table 7 lists the expected impact of price collusion on welfare, ignoring product collusion.

Table 7
Welfare Effects Of Price Collusion

	Welfare Effect (in millions)
Δ Consumer Surplus	-\$17.143
Δ Producer Surplus	\$2.872
Δ Social Surplus	-\$14.271

The antitrust authorities would likely infer large effects on consumer and social welfare from these estimates. Extrapolating across all supermarkets suggests that \$237.368 million was lost in social surplus in 2013. However, this estimate is misleading for two reasons. First, it does not account for how firms colluding on products might reposition their flavor offerings if prevented from colluding on price. Therefore, this estimate would not necessarily reflect the welfare gains from preventing price collusion. Second, it does not account for the direct effect of product collusion on welfare which is ambiguous. To address these issues, I now consider the effect of product space collusion on welfare.

8.2 Effect Of Eliminating Product & Price Collusion

To account for the effect of product collusion on welfare, I compute two additional counterfactuals in which I allow firms to reoptimize both their product and pricing decisions. In the first, I set both the product and price collusion parameters to zero and find the equilibrium set of products and prices when the firms are not allowed to collude in either stage. This allows me to measure the direct effect of product and product collusion on welfare. In the second, I hold the product collusion parameters fixed at the levels in Table 4 and set the price collusion parameters to 0. It is unknown if the firms would collude to the same degree on the product space if they could not collude in price, making this second counterfactual highly speculative. I therefore present this counterfactual merely to highlight a potential problem associated with implementing antitrust policies without considering the full scope of collusion: namely the welfare effects of preventing price collusion are ambiguous when firms also collude on the product space.

To find the equilibrium set of products in both counterfactuals, I employ an algorithm developed in Fan and Yang (2016), which the authors describe in depth. I will provide a brief summary here. I define the product space as consisting of all non-seasonal flavors offered in my sample. The algorithm begins with both firms offering the subset of products observed in the data. I then allow Ben & Jerry's to consider all one product deviations. These include all one-product additions, one-product removals, and one-product switches. At every deviation, I update the firm's prices to satisfy the pricing first order conditions. If Ben & Jerry's does not find it profitable to deviate,[46] then the observed set of products is Ben & Jerry's best response and I consider all one-product deviations for Häagen-Dazs. If Ben & Jerry's finds it profitable to deviate, I consider all one-product deviations from the set of products that yeilded the highest payoff, continuing in this way until Ben & Jerry's does not find it profitable to deviate. The algorithm continues until neither firm can benefit from any one-product deviation given the actions of its rival.

Because both the set of products is large and the firms update their prices at each deviation, it is not feasible to perform the counterfactual in all markets. Thus, I focus attention on the median market based on the average profitability of the firms. In this supermarket-month, the total expected profits, excluding fixed costs, were \$399.25 and \$412.34 for Ben & Jerry's and Häagen-Dazs respectively. The firms jointly offered 66 products in this supermarket-month, 37 Ben & Jerry's flavors and 29 Häagen-Dazs flavors. A summary of the characteristics in these flavor sets is reported in the first column of Table 8.

Table 8
Characteristics Of Flavor Sets

	Observed Product Set		No Price or Prod. Collusion		No Price Collusion Only	
	Ben & Jerry's	Häagen-Dazs	Ben & Jerry's	Häagen-Dazs	Ben & Jerry's	Häagen-Dazs
Number Flavors	37	29	42	28	15	23
Vanilla	0.270	0.241	0.238	0.250	0	0.348
Chocolate	0.243	0.172	0.262	0.214	0.400	0.182
Coffee	0.054	0.069	0.071	0.036	0	0.046
Fruit	0.162	0.138	0.190	0.179	0.333	0.091
Mixins	2.432	1.345	2.357	1.143	2.133	1.091

I measure the overall welfare effects from both product and price collusion in this selected market. The first column of Table 9 replicates the analysis in Section 8.1, comparing the observed collusive outcome to the counterfactual in which the products are held fixed and the price parameters are set to 0. The second column takes into account both product and price collusion, comparing observed outcomes to the counterfactual in which all collusion parameters are set to 0. I find that failure to account for product space collusion considerably underestimates the negative effect of collusion on consumer and social surplus.

[46] Profitability has different meanings in the two counterfactuals. In the first counterfactual, the payoff to firm b is just its own profits, π_b. In the second, where the firms are allowed to collude on the product space, the payoff to each firm b is measured as $\pi_b + \theta_{1,b}\pi_{-b}$.

Table 9
Effect Of Collusion In Median Market

	When Accounting for the Presence of . . .	
	Price Collusion Only	Prod. & Price Collusion
Δ Consumer Surplus	-\$236.27	-\$281.24
Δ Total π	\$40.75	\$34.77
Δ Fixed Cost	\$0	-\$31.53
Δ Social Surplus	-\$195.52	-\$278.00

I do find that failure to account for product space collusion overstates the positive effect on producer surplus. This is because the counterfactual in which the product set is held fixed and the price collusion parameters set to 0 does not reflect the optimal product and pricing choice that the firms would make at those parameter values. In this market, I find that the producer surplus would be higher if the firms did not collude in either stage than if they were forced to offer the produced products when they could no longer collude on price.

Given that firms would have chosen different product offerings in this market if the antitrust authorities had prevented price collusion, it is important to perform the second counterfactual as an illustration even if it is subject to the Lucas Critique. In doing so, I find that the firms would have dramatically curtailed the set of flavors offered if they could not collude in price but continued to collude to the same degree on the product space. In particular, I find two equilibria, one in which Ben & Jerry's and Häagen Dazs offer 15 and 23 products respectively, the other in which they offer 24 and 20 flavors. The third column of Table 8 summarizes the product sets for the first equilibrium. Strikingly, not only does the number of products shrink, but the differentiation in the offerings between the brands increases a great deal. As a result, I find that consumers are better off when firms are colluding both in prices and products than when firms can only collude on the product space. In particular, consumer surplus falls by \$195.52 in the first equilibrium when price collusion is prevented and \$45.00 in the second equilibrium. This suggests a potential caveat if regulators focus only on price collusion.

Taken together, these findings are fairly intuitive. When the firms can only collude in product choice, they withdraw considerably from their rival's designated part of the product space. This allows them to avoid significant price competition. If the firms have the ability to also collude in prices, they are able to move their product sets closer to each other without triggering price competition. Finally, when the firms cannot collude in either stage, the firms directly enter each other's space. In my example, as long as the firms are colluding on the choice of products, Ben & Jerry's does not offer chocolate ice cream. However, if the firms were prevented from colluding at all, chocolate would be sold by both firms.

It is important to bear in mind that these results are derived from only one supermarket month, albeit one that has been chosen to be representative. Because the theoretical predictions are ambiguous, it is possible that product space collusion could have the opposite effect on welfare in other settings. I plan on checking the robustness of these results in future work. However, they do suggest the importance of accounting for product collusion.

9 Conclusion

In this paper, I presented a methodology by which researchers can measure product space collusion. I then applied that method to the market for super-premium ice cream during 2013. I found substantial evidence that Ben & Jerry's and Häagen-Dazs colluded on both the set of products offered and the prices charged for these products. I also constructed counterfactuals to quantify the effects of product and price collusion on firm actions and welfare.

There are four main implications of my paper for antitrust policy. First, ignoring product collusion can result in either anticompetitive behavior going undetected or the welfare effects being miscalculated. Second, there exists the possibility that policy interventions to prevent price collusion may have negative impacts on welfare when product collusion is not considered. Third, collusion does not necessarily reduce welfare when both price and products are considered, so it should be evaluated on a case-by-case basis rather than being per se illegal. Finally, in merger analysis, the antitrust authorities often allow firms to merge if they do not compete in the same markets. For example, airlines are often permitted to merge so long as they do not offer service on the same point-to-point routes. It is possible though that airlines service different routes because of product collusion and that competition would have caused them to compete directly. Thus, it seems that current antitrust policy may be rewarding firms for behaving anticompetitively with respect to their product choices.

I would like to extend this work in two directions. First, this paper only addresses the static implications of product space collusion, yet there may also exist dynamic effects. Beginning with Schmalensee (1978), a strain of the IO literature has investigated how incumbents in a market for differentiated products could deter entry by packing the product space. If firms produce a wide array of differentiated products, space may not exist for potential entrants. The introduction of new products often comes with a fixed R&D cost. In a oligopolistic market, firms which introduce new products pay the entire R&D cost but share the entry-deterring benefits with all incumbents. Thus, a free riding problem might exist which, from the private perspective of the firms, prevents optimal product packing. However, if firms collude on the product space by partitioning it amongst themselves, then each firm has the incentive to optimally pack its part of the product space to deter entry.

Second, the literature has long been puzzled by the phenomenon of uniform pricing. In working on this paper, I have been surprised by the small effect that changing the product set has on brand prices. Thus, it is possible that uniform brand prices serve to stabilize product collusion by reducing the profitability of defecting in the product space. Consider Ben & Jerry's replacing a flavor with chocolate. If Häagen-Dazs priced all its flavors individually, this defection by Ben & Jerry's would likely have a large effect on the price of Häagen-Dazs chocolate. However, the uniform brand price insulates Häagen-Dazs to a degree.

References

Ackerberg, Daniel. 2016. "Timing Assumptions and Effciency: Empirical Evidence in a Production Function Context." working paper.

Ackerberg, Daniel, and Jinyonh Hahn. 2015. "Some Non-Parametric Identification Results using Timing and Information Set Assumptions." working paper.

Andrews, Donald, and Gustavo Soares. 2010. "Inference for Parameters Defined by Moment Inequalities Using Generalized Moment Selection." *Econometrica*, 78(1): 119–157.

Berry, Steven, and Philip Haile. 2014. "Identification in Differentiated Products Markets Using Market Level Data." *Econometrica*, 82(5): 1749–1797.

Berry, Steven, James Levinsohn, and Ariel Pakes. 1995. "Automobile Prices in Market Equilibrium." *Econometrica*, 63(4): 841–890.

Black, Michael, Gregory Crawford, Shihua Lu, and Halbert White. 2004. "A Virtual Stakes Approach to Measuring Competition in Product Markets." working paper.

Bresnahan, Timothy. 1982. "The Oligopoly Solution Concept is Identified." *Economics Letters*, 10: 87–92.

Bresnahan, Timothy. 1987. "Competition and Collusion in the American Automobile Industry: The 1955 Price War." *The Journal of Industrial Economics*, 35(4): 457–482.

Corts, Kenneth. 1999. "Conduct Parameters and the Measurement of Market Power." *Journal of Econometrics*, 88(8): 227–250.

Crawford, Gregory, and Ali Yurukoglu. 2012. "The Welfare Effects of Bundling in Multichannel Television Markets." *American Economic Review*, 102(2): 643–685.

Crawford, Gregory, and Matthew Shum. 2007. "Monopoly Quality Degradation and Regulation in Cable Television." *Journal of Law and Economics*, 50(1): 181–209.

Draganska, Michaela, Michael Mazzeo, and Katja Seim. 2009. "Beyond Plain Vanilla: Modeling Joint Product Assortment and Pricing Decisions." *Quantative Marketing and Economics*, 7(2): 105–146.

Dube, Jean-Pierre, Jeremy Fox, and Che-Lin Su. 2012. "Improving the Numerical Performance of Static and Dynamic Aggregate Discrete Choice Random Coefficients Demand Estimation." *Econometrica*, 80(5): 2231–2267.

Fan, Ying. 2013. "Ownership Consolidation and Product Characteristics: A Study of the US Daily Newspaper Market." *American Economic Review*, 103(5): 1598–1628.

Fan, Ying, and Chenyue Yang. 2016. "Competition, Product Proliferation and Welfare: A Study of the U.S. Smartphone Market." working paper.

Fershtman, Chaim, and Ariel Pakes. 2000. "A Dynamic Oligopoly with Collusion and Price Wars." *RAND Journal of Economics*, 31(2): 207–236.

Gandhi, Amit, and Jean-Francois Houde. 2015. "Measuring Substitution Patterns in Differentiated Products Industries." working paper.

Gandhi, Amit, Zhentong Lu, and Xiaoxia Shi. 2013. "Estimating Demand for Differentiated Products with Error in Market Shares." working paper.

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

Hackner, Jonas. 1995. "Endogenous Product Design in an Infinitely Repeated Game." *International Journal of Industrial Organization*, 13(2): 277–299.

Hendel, Igal, and Aviv Nevo. 2006*a*. "Measuring the Implications of Sales and Consumer Inventory Behavior." *Econometrica*, 74(6): 1637–1673.

Hendel, Igal, and Aviv Nevo. 2006*b*. "Sales and Consumer Inventory." *RAND Journal of Economics*, 37(3): 543–561.

Mailath, George, Volker Nocke, and Lucy White. forthcoming. "When and How the Punishment Must Fit the Crime." *International Economic Review.*

Mas-Colell, Andreu, Michael Whinston, and Jerry Green. 1995. *Microeconomic Theory.* New York, NY:Oxford University Press.

Mazzeo, Michael. 2002. "Product Choice and Oligopoly Market Structure." *RAND Journal of Economics*, 33(2): 221–242.

Miller, Nathan, and Matthew Weinberg. forthcoming. "Understanding the Price Effects of the MillerCoors Joint Venture." *Econometrica.*

Nevo, Aviv. 1998. "Identity of the Oligopoly Solution Concept in a Differentiated Products Industry." *Economics Letters*, 59(3): 391–395.

Nevo, Aviv. 2001. "Measuring Market Power in the Ready-to-Eat Cereal Industry." *Econometrica*, 69(2): 307–342.

Orhun, A. Yesim. 2013. "Spatial Differentiation in the Supermarket Industry: The Role of Common Information." *Quantitive Marketing and Economics*, 11(1): 3–37.

Orhun, A. Yesim, Sriram Venkataraman, and Orafeeo Chintagunta. 2015. "Impact of Competition on Product Decisions: Movie Choices of Exhibitors." *Journal of Marketing Science*, 35(1): 73–92.

Pakes, Ariel. 2010. "Alternative Models for Moment Inequalities." *Econometrica*, 78(6): 1783–1822.

Pakes, Ariel, Jack Porter, Kate Ho, and Joy Ishii. 2011. "Moment Inequalities and Their Application." working paper.

Rotemberg, Julio, and Garth Saloner. 1986. "A Supergame-Theoretic Model of Price Wars During Booms." *American Economic Review*, 76(3): 390–407.

Schmalensee, Richard. 1978. "Entry Deterrence in the Ready-to-Eat Breakfast Cereal Industry." *Bell Journal of Economics*, 9(2): 305–327.

Small, Kenneth, and Harvey Rosen. 1981. "Applied Welfare Economics with Discrete Choice Models." *Econometrica*, 49(1): 105–130.

Sullivan, Christopher. 2017. "Will You Just Hold Still?: Using Cross- Sectional Variation to Circumvent the Corts Critique." working paper.

Sweeting, Andrew. 2013. "Dynamic Product Repositioning in Differentiated Product Markets: The Effect of Fees for Musical Performance Rights on the Commercial Radio Industry." *Econometrica*, 81(5): 1763–1803.

Wollmann, Thomas. 2016. "Trucks without Bailouts: Equilibrium Product Characteristics for Commercial Vehicles." working paper.

Xu, Xu, and Kalyn Coatney. 2015. "Product Market Segmentation and Output Collusion Within Substitute Products." *Journal of Economics and Business*, 77: 1–15.