Common Ownership and Competition in the Ready-To-Eat Cereal Industry*

PRELIMINARY AND INCOMPLETE
PLEASE DO NOT CITE OR CIRCULATE
(NOTE: THIS VERSION HAS BEEN REDACTED DUE TO OUR DATA LICENSE AGREEMENT)

Matthew Backus† Christopher Conlon‡ Michael Sinkinson§

September 6, 2018

Abstract

Publicly-traded firms have a fiduciary duty to shareholders, which motivates the assumption of firm-level profit maximization. However, if those shareholders have an interest in competitors, the assumption might be misguided. We begin by documenting a broad increase in common ownership over time and show the implications for firm incentives through the lens of profit weights. We then investigate this concern in the ready-to-eat cereal industry, using matched product-market and ownership data. We estimate a structural model of demand and pricing that nests alternative objective functions in order to test alternative models. Finally, we quantify the implied effect of common ownership on prices.

*Thanks to... All remaining errors are our own.
†Columbia University and NBER, matthew.backus@columbia.edu
‡New York University, cconlon@stern.nyu.edu
§Yale SOM and NBER, michael.sinkinson@yale.edu

1 Introduction

Publicly traded corporations, starting with the Dutch East India Company, have been the backbone of the global economy for hundreds of years. The fundamental promise of the corporation is that it maximizes shareholder value. The board of directors, with a fiduciary duty to represent the interests of shareholders, designs incentives and delegates decisions to management. The goal of maximizing shareholder value aligns closely with that of maximizing profits, as shares are valued at the discounted present value of all future earnings.

More recently, concerns have been raised that through the rise of large index funds, ownership of public corporations has become concentrated among a small number of investors, and that this may distort the decisions firms make. In particular, the two largest asset managers, Blackrock and Vanguard, have a combined assets under management total of $10.2 trillion.[1] To put that number in context, the total market capitalization of the S&P 500 on December 31, 2017 was $23.9 trillion. This is concerning because the classic model of the corporation implicitly assumes shares are held by undiversified investors. Suppose a firm could take an action that would cause its own share price to rise, but that of a rival to fall, e.g. by obtaining some competitive advantage. In the classic model, the firm would take such an action, narrowly pursuing own profits. However, if the firm's shareholders also held shares in the rival, the firms incentives are less clear.

This issue has gained a great deal of attention in the past several years, with a number of attempts to measure the empirical impact of common ownership on firm pricing (Azar et al., 2016, 2017). These papers have argued that airfare and bank interest rates and fees are responsive to the level of common ownership in a market. The approach in these papers is to run reduced-form regressions of prices on concentration measures such as HHI and ownership measures such as MHHI. The challenge with measurement here, as pointed out by O'Brien (2017) , is that a regression of prices on MHHI is not motivated by theory. As a result, some have challenged (e.g. Kennedy et al. (2017)) the empirical findings of an effect of common ownership.

The theory of common ownership implies that firms should place weights on rival firms' profits when making strategic choices. Our analysis of this hypothesis proceeds in two parts. First, we gather data on instutitional ownership for all firms in the S&P 500 for the

[1]These figures are dated September 30, 2017. Blackrock reports $5.7 trillion AUM and Vanguard reports $4.5 trillion AUM.

period 1980-2016. We use this data to compute the implied *profit weights* that each firm places on each other firm's profits. This allows us to characterize trends and heterogeneity in common ownership. The standard O'Brien and Salop (2000) model, which arises from a voting model of corporate governance, yields a starkly linear increasing trend for profit weights with weights tripling from 0.2 to 0.6 between 1980 and the present day. Independent of the model of corporate governance, however, we also find that the weights are positively correlated with both market capitalization as well as the ownership share of retail investors.

In order to consider welfare effects, we need to choose a product market and a strategic setting for common ownership. Therefore, in the next part of the paper we explore how common ownership would manifest itself in a typical Industrial Organization estimation framework. We show that the standard firm profit maximization approach can be adapted to common ownership by including measures of common ownership in the "ownership matrix" used by the firm to set optimal prices. In the spirit of Bresnahan (1987) and Nevo (2001), we consider several specifications for how common ownership would enter into the ownership matrix.

We apply our approach to the market for ready-to-eat cereals, using scanner-level data over a ten year time horizon. The chosen industry is an ideal one for this question as it exhibits variation in ownership concentration across firms, as well as over time, in addition to a high level of concentration to begin with (C_4 is approximately 85%). For example, The Kellogg Foundation is a major (undiversified) shareholder of Kelloggs, Inc, although its share has fallen over time. We also observe that the products sold under the Post brand changed hands several times, from being a part of Kraft foods, to being a part of Ralcorp, and finally to being an independent firm, all of which affect the extent to which its shares are commonly owned.

In this context we illustrate the economic content of different conduct assumptions on firm price-setting in ready-to-eat cereal: firm-level profit maximization, total collusion, and variants of the common ownership hypothesis. We construct counterfactual predictions of prices with and without common ownership in order to quantify the economic significance of the hypothesis, and we endeavor to test the two models of conduct using exclusion restrictions on marginal costs.

This paper contributes to the evolving literature on common ownership in several ways. First, we document the implied profit weights each S&P 500 firm would place on every other firm's profits under different common ownership assumptions. Second, we provide a framework for including common ownership effects in structural empirical research on differentiated product markets. Third, we assess the impact of common ownership in a well-studied product market and quantify the potential impact of different forms of common ownership.

2 Related Literature

Empirical testing of firm conduct is an endeavor with a storied history in industrial organization. Modern methods build on advances in structural modeling of demand and firm pricing. From a methodological standpoint, Berry and Haile (2014) iterate on the insight of Bresnahan (1982) to show that models of conduct are testable if demand is identified subject to restrictions on the cost structure. Examples of this approach include Bresnahan (1987), Nevo (2001), and Miller and Weinberg (2017).[2] With respect to this literature, we pose the common ownership hypothesis as an alternative model of conduct, and employ some of the same identification results that have been used to test for collusion and portfolio pricing.

The theoretical foundations of the common ownership hypothesis are not new. (Rotemberg, 1984) offered the earliest model in which diversification of shareholders affects the character of imperfect competition in product markets. Bresnahan and Salop (1986) and O'Brien and Salop (2000) treats the same question, characterizing "partial mergers," where diversified ownership by shareholders and cross-ownership among firms effects degrees of control. In particular, Bresnahan and Salop (1986) introduced the modified HHI (MHHI) concentration measure to capture such partial control.

The recent revival of interest in the common ownership hypothesis follows from a handful of empirical studies that appear to show large effects on pricing in product markets. The main two instigating endeavors are Azar et al. (2016), which studied the effects of common ownership on bank fees, and Azar et al. (2017), which uses the Blackrock acquisition of Barclays as an instrument to study the effect of common ownership on airfares. Neither of the above papers address the question of finding a mechanism – precisely *how* do large common investors affect prices? Anton et al. (2016) propose that it is through the flattening of executive compensation schemes with respect to firm performance. They exploit scandal-

[2] Nevo (2001) also studies ready-to-eat cereal to test models of conduct and is the most proximate point of departure for our endeavor. We iterate on that seminal work by using an updated sample (1990-1996 vs 2000 - 2015), more stores (6 versus thousands), more detailed characteristics data, store-specific measurement of market size and, most importantly, an alternative conduct hypothesis.

induced changes in common ownership and show that they are correlated with attributes of corporate compensation in a way that is consistent with reduced competition in product markets.

This literature has not been without criticism, in particular for its reduced-form approach. The explanatory variables of interest are concentration indices such as MHHI, an approach to studying markets that has fallen out of favor in industrial organization. Besides the obvious concerns for identification (which may be addressed with instruments), these concentration measures have no monotonic relationship to the main outcome of interest – product market prices – as O'Brien (2017) reminds us. Notwithstanding these concerns about the empirical results, this burgeoning literature has attracted the interest of the legal community (Elhauge, 2016) and antitrust authorities, and there is already interest in regulatory remedies (Posner et al., 2017). The literature is further reviewed in a related paper by the authors (Backus et al. (2018)). It is a pressing concern then, to develop a structural approach that can guide antitrust and regulatory authorities in evaluating both the problem and potential remedies. To our knowledge, the only competing paper that endeavors a structural analysis of the common ownership hypothesis is Kennedy et al. (2017), which takes a different approach and studies the airline industry.[3]

3 Common Ownership: Theoretical Preliminaries

We begin with a generic setup: a firm making a strategic choice x_f and whose profits are given by $\pi_f(x_f, x_{-f})$ and thus depend on their rivals' choice as well. Under the maintained hypothesis of firm profit maximization, the profit function constitutes the objective function of the firm, and it is in this framework that economists have traditionally modeled behavior ranging from pricing to R&D. This is occasionally motivated by the claim that the firm answers to its investors, who should be unwilling to provide capital should the firm fail to at least mimic profit maximization Friedman (1953).

So consider the payoffs of an investor – for our purposes, a shareholder. That shareholder

[3]The papers are quite different in approach and, given the gravity of the question, we view this heterogeneity in method as complementary. In particular they study airlines, mirroring Azar et al. (2017), while we have deliberately chosen an industry with simpler pricing practices; they estimate a nested logit model of demand while ours is a random coefficient logit; and finally they estimate a conduct parameter where we follow the "menu" approach of Nevo (2000) and study discrete, fully specified, conduct models.

may be invested in several firms. We assume that shareholder s has cash-flow rights equal to the fraction of the firm f they own, denoted β_{fs}, and therefore the profit of the shareholder is given by $\pi_s = \sum_{\forall g} \beta_{gs}\pi_g$.

However shareholders portfolios differ, and so then do their profit functions. If we conceive of firms as mechanisms for maximizing shareholder value, how are these potentially divergent interests aggregated into a choice function? The most conventional approach, which builds directly on Rotemberg (1984), is due to O'Brien and Salop (2000). In this framework, the firm acts to maximize the profits of shareholders, subject to a vector of Pareto weights. Letting Q_f denote the proposed objective function of the firm, γ_{fs} denote the weight that firm f places on shareholder s, and β_{fs} the cash flow rights shareholder s in firm f, then

$$\begin{aligned} Q_f(x_f, x_{-f}) &= \sum_s \gamma_{fs} \cdot \pi_s(x_f, x_{-f}) = \sum_s \gamma_{fs} \cdot \left(\sum_{\forall g} \beta_{gs} \cdot \pi_g(x_f, x_{-f}) \right) \\ &= \sum_s \gamma_{fs}\beta_{fs}\pi_f + \sum_s \gamma_{fs} \sum_{\forall f \neq g} \beta_{gs} \cdot \pi_g \\ &\propto \pi_f + \sum_{g \neq f} \underbrace{\frac{\sum_s \gamma_{fs}\beta_{gs}}{\sum_s \gamma_{fs}\beta_{fs}}}_{\equiv \kappa_{fg}^{OS}(\gamma_f, \beta)} \pi_g = \pi_f + \sum_{g \neq f} \kappa_{fg}^{OS}(\gamma_f, \beta)\pi_g. \end{aligned} \tag{1}$$

In the last line we show that we can rewrite the maximization of shareholder profits into a maximization problem over own- and competing firms' profits. In our notation, κ_{ff} is always normalized to one, so that κ_{fg} can be interpreted as the relative value of a dollar of profits accruing to firm g in firm f's maximization problem. This may be greater than zero if firm fs shareholders also have cash flow rights in firm g. We refer to these κ terms going forward as *profit weights*.

An alternative specification is offered in Crawford et al. (2018).[4] There, the firm aggregates shareholder preferences by taking a γ-weighted convex combination of their desired weights. Define

$$\tilde{\beta}_{fs} = \frac{\beta_{fs}}{\sum_g \beta_{gs}}.$$

[4] It is important to note that Crawford et al. (2018) are not considering the common ownership hypothesis directly but rather examining incentives for vertical integration and bargaining among MVPDs and content providers where the former often have a partial ownership stake in the later.

Then the vector $\tilde{\beta}_s$ is the (non-normalized) ideal profit weights for shareholder s. The firm maximizes

$$\begin{aligned} Q_f(x_f, x_{-f}) &= \sum_s \gamma_{fs} \sum_g \tilde{\beta}_{gs} \cdot \pi_g \\ &\propto \pi_f + \sum_{g \neq f} \underbrace{\frac{\sum_s \gamma_{fs} \tilde{\beta}_{gs}}{\sum_s \gamma_{fs} \tilde{\beta}_{fs}}}_{\equiv \kappa_{fg}^{CLWY}(\gamma_f, \beta)} \pi_g = \pi_f + \sum_{g \neq f} \kappa_{fg}^{CLWY}(\gamma_f, \beta) \pi_g. \end{aligned} \tag{2}$$

Comparing the OS and CLWY weights, the latter will tend to put relatively less emphasis on large, diversified investors. Fixing γ_{fs}, an investor's influence on the matrix κ^{OS} is in proportion to their total holdings. This is intuitive – since it values the marginal dollar of shareholder s at γ_{fs}, the larger their holding the larger their exposure, and therefore the greater their effect on κ^{OS}. This channel is explicitly muted in the CLWY weights, which normalizes by the scale of the shareholder's total position. Rather than a maximization of surplus, the economic intuition here corresponds to a proportional representation voting model.

This paper will focus on the $kappa^{OS}$ formulation going forward, although we will explore alternative formulations in the Appedix.

If we summarize κ values in a square matrix, where the diagonal is always one and the f,gth element is the weight f places on the profits of the gth firm in their maximization problem, these are cooperation matrices. The cooperation matrix has a long history in the broader IO literature going back at least as far as Bresnahan (1987). Here we have defined it at the level of the firm, but we could have defined it at the level of an individual product. Following Bresnahan's original work, absent collusive behavior, κ^{MAX} is simply the identity matrix. Alternatively, under collusion, κ^{COL} is everywhere equal to one. Cooperation matrices are also used to model mergers, as in Nevo (2000): pre-merger $\kappa_{fg} = \kappa_{gf} = 0$ whereas post-merger $\kappa_{fg} = \kappa_{gf} = 1$. Miller and Weinberg (2017) consider tacit collusion in the beer industry and actually attempt to estimate κ_{fg} for the participants. Similarly, Kennedy et al. (2017) estimate a parameter τ in front of all off-diagonal κ_{fg} in order to test the common ownership hypothesis in airlines.

The cooperation matrix κ is a functions of γs and βs, so where do these objects come from? For most publicly traded firms in the US, the cash flow rights of shareholder s in firm f are given by the fraction they own of total shares outstanding. This is measurable for large institutional investors.

The second element, the weights the firm places on its shareholders, is less transparent. We follow the literature in assuming that $\gamma_{fs} = \beta_{fs}$.[5] This reflects the "one share one vote" rule that characterizes most publicly traded firms in the US. Any formulation of the γs is implicitly a model of corporate governance, and one on which the theory offers precious little guidance. Happily, ours is monotonic increasing and continuous in holdings, two attractive, if inconclusive, features.[6] Observe also that the only way to nest the standard model of own-profit maximization in this framework is to assume that the firm places zero weight on any shareholders with diversified portfolios.

Example 1: Consider a market with three firms. Firm 1 is privately held, in its entirety, by an undiversified investor. Firms 2 and 3 have the following identical ownership structure: 60 percent of each is held by small, undiversified retail investors. 20 percent of each are held, respectively, by two large, undiversified investors. The final 20 percent of each is held by a single, diversified investor. This ownership patten is summarized in Table 1.

Table 1: Example 1 Ownership Structure

	Firm 1	Firm 2	Firm 3
Investor 1	100%	-	-
Investor 2	-	20%	-
Investor 3	-	-	20%
Investor 4	-	20%	20%
Retail Share	-	60%	60%

This yields the following set of profit weights:

[5]When $\gamma_{fs} = \beta_{fs}$, then for the construction of the OS weights it is innocuous to ignore small retail investors (because, for these investors, β_{fs}^2 is vanishingly small). This is not the case for the alternative formulations such as CLWY weights. Here we would have to make the additional assumption that $\gamma = 0$ for retail investors, reflecting the low frequency with which they exercise their voting rights.

[6]As an example where these features may fail, consider κ^{OS} in the case where $\gamma = 1$ for all shareholders of firm f, i.e. the firm maximizes their shareholder's portfolio value. This model introduces a potentially large discontinuity with the purchase of the first share.

$$\kappa = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 1/2 \\ 0 & 1/2 & 1 \end{bmatrix}$$

To see how this calculation is done, denote column j of Table 1 as C_j (excluding the bottom row, the retail share which is assumed to have no weight). Then, the profit weight firm f has on firm g's profit is $\kappa_{fg} = (C_f{}' \cdot C_g)/(C_f{}' \cdot C_f)$. This example highlights that the profit weights can be quite large.

Example 2

Now consider an alternative market with just two firms. The vast majority of both firms are held by a large set of undiversified retail investors. A boundedly small fraction of both firms is held by a finite set of N symmetric, diversified investors who each hold 1 percent of firm one and x percent of firm two. This ownership pattern is summarized in Table 2.

Table 2: Example 2 Ownership Structure

	Firm 1	Firm 2
Investor 1	1%	x%
Investor 2	1%	x%
⋮	⋮	⋮
Investor N	1%	x%
Retail Share	(100-N) %	(100-Nx)%

Then, we would have the following κ matrix of profit weights:

$$\kappa = \begin{bmatrix} 1 & x \\ 1/x & 1 \end{bmatrix}$$

The calculation follows in the same manner as Example 1. This example highlights a few points about profit weights. First, letting $x = 1$, we have that an arbitrarily small share of ownership has effected full collusion. Note that the profit weights do not depend on N. This comes about because the large share of retail investment has effectively diluted control rights. If, say, 90 percent of a firm is held by retail investors who are uninvolved in corporate governance, then the remaining ten percent has ten-to-one control rights, the same kind of dilution of control rights that leads to distortions of incentives in the literature on tunneling. Second, letting $x \neq 1$, we see that it is entirely possible for a firm to place weight greater than one on the profits of another firm, another feature that echoes tunneling.

The reader will note that we have omitted any discussion of incentives or mechanisms. Do large institutional investors, e.g., Blackrock and Vanguard, who make up much of the driving variation have an incentive to engage in corporate governance to increase the value of their customers' portfolios? What are the mechanisms by which their preferences are communicated to policy-making in the firms in which they have invested? These forensic questions are justifiably contentious, and we do not have new answers. The interested reader can find a plethora of motivating anecdotes on these points in prior work. Here we are interested in understanding the economic implications of the hypothesis, should both the will and the means exist, and developing econometric methods to test such a hypothesis against the standard model of own-firm profit maximization.

4 Trends and Patterns in Common Ownership

While there is broad agreement that common ownership is on the rise – under the premise that there is growing concentration among highly diversified institutional investors – little is known about the magnitude of the trend or patterns therein. What sectors are most affected? What types of firms seem most exposed to common ownership? And, what is drives that heterogeneity?

In this section we take a broad, descriptive approach to answering these questions, with an eye to motivating the industry-based approach in Sections 5 through 9. We compute common ownership weights (κ's) among all firms in the S&P 500 for the period 1980–2016, excluding a relatively small set of 49 firms that use dual-class shares to separate control rights from cash-flow rights. We use the S&P 500 as it is considered to be representative of the US economy and many investment funds offer products tied to the constituent firms in one way or another.

Our work parallels a complementary effort to compute MHHI indices by two-digit sector in Anton et al. (2016). We prefer direct measurement of κ for two reasons. First, computation of MHHI requires market definition and market shares. Besides the difficulty of market definition, this raises problems of data quality when many firms are private or foreign-owned. Trends in either feature will introduce biases in the estimate time path of MHHI. Second, MHHI is only interpretable in Cournot oligopoly, and this is only a reasonable approximation of a very small slice of the economy. Common ownership weights, in contrast, are well-defined for any pair of firms because they make no assumption on the nature of the strategic interaction. For this same reason it is important to emphasize that this exercise does not have welfare implications. While much of the recent literature on common ownership has emphasized anticompetitive effects on pricing of substitute goods, a subject to which we turn soon, there is no such presumption here. If two firm are entirely unrelated, then the common ownership weight between them may have no effect on behavior. Alternatively, if they are vertically related – e.g., a manufacturer and a supplier – then common ownership may be welfare-enhancing if it mitigates double-marginalization. And, as López and Vives (2018) show, when the strategic variable is something other than prices, in their case R&D, welfare effects could go either way.

4.1 Data on Common Ownership

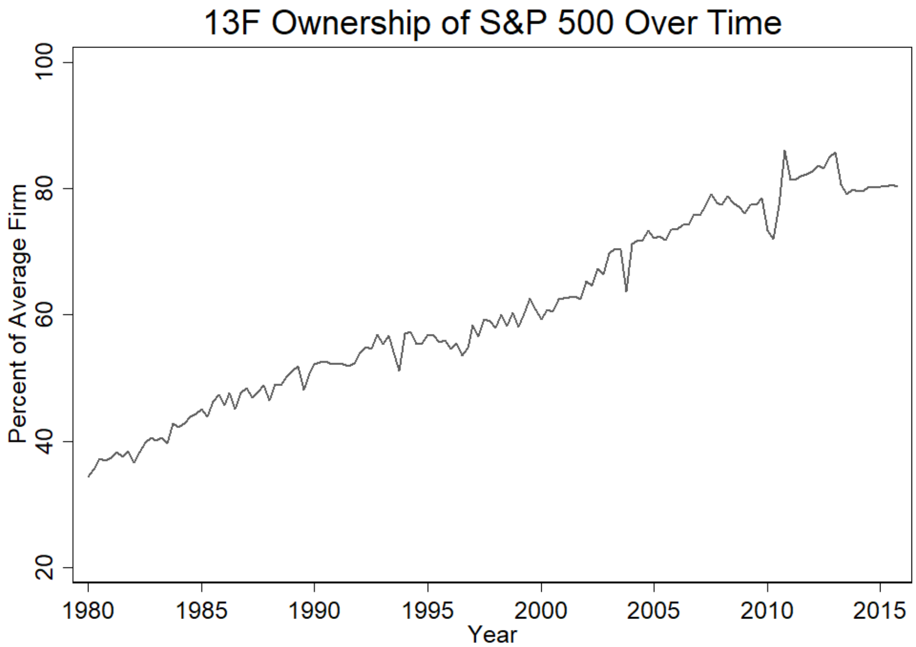
The primary data source for common ownership is Thomson Reuters, who consolidate the "13f" filings required by the Securities and Exchange Commission (SEC) for all investment managers with over $\$100 MUSD$ in holdings of a list of "13f" securities. These data go back to 1980. The filings are quarterly and mandatory. These data are available to researchers through Wharton Research Data Services (WRDS).

Recently, WRDS and some researchers (Ben-David et al. (2018)) noticed data quality issues regarding the Thomson dataset, and they have worked to resolve these issues. We use the July 2018 update provided by WRDS below. We consolidate managers that report multiple subsidiaries. Data quality issues are discussed in more depth in Backus et al. (2018).

Data on the constituents of the S&P 500 come from Compustat. We also gather data on shares outstanding from the Center for Research in Security Prices (CRSP). In terms of the raw data, Figure 1 shows that 13f investment managers own an increasing share of the S&P 500 over time, rising from below 40% to over 80%.

Figure 2 highlights holdings by some large investment managers that have received attention in the common ownership literature: Blackrock, Vanguard, State Street, and Fidelity. The

Figure 1: Share of S&P 500 Owned by 13f Managers



plot shows that these firms each currently hold between 3 and 7% of a typical S&P 500 firm, and that this has increased over time.

4.2 Profit Weights

Figure A-4 plots the average κ for every pair of S&P 500 firms by year. Using the O'Brien and Salop (2000) formulation, we see a stark linear trend, growing from an average of 0.2 to 0.6 between 1980 and 2016. We explore alternative κ formulations in Appendix A.1. We interpret the rise in the 80's, 90's, and early aughts to the rise of 401(k) savings plans, which emerged in the 80's and grew to become the primary investment vehicle for households, also highly diversified.

Next we turn to understanding heterogeneity in common ownership weights among our sample of S&P 500 firms. In Figure 5 we plot the profit weights against log market capitalization as well as the retail share of investors, for twenty equal-sized bins. [7]. All plots absorb year fixed effects to account for levels of average nominal capitalization or retail share of aggregate

[7] Results for CLWY weights are similar and available in Appendix A.1

Figure 2: Share of Typical Firm Owned

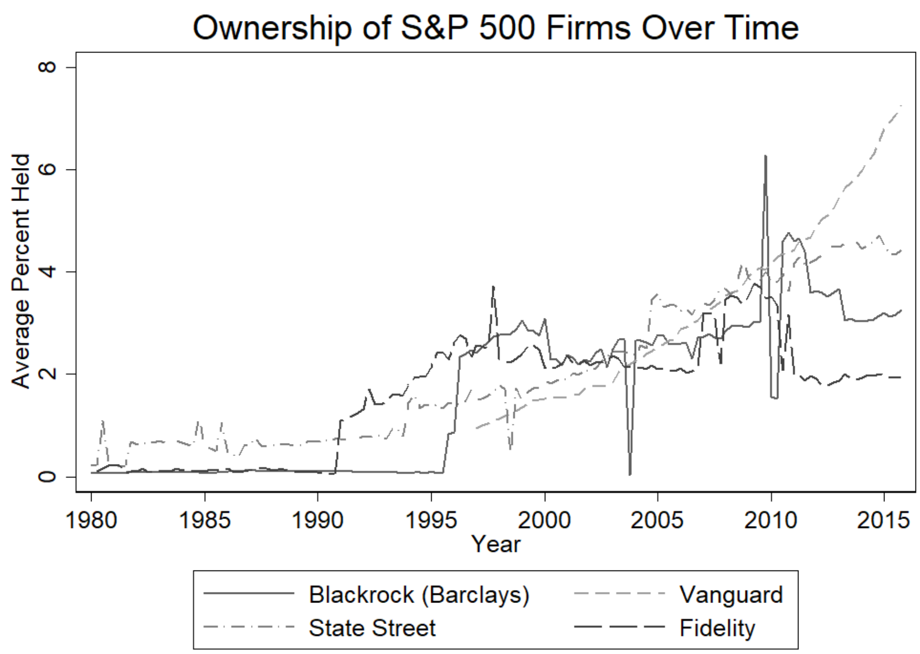
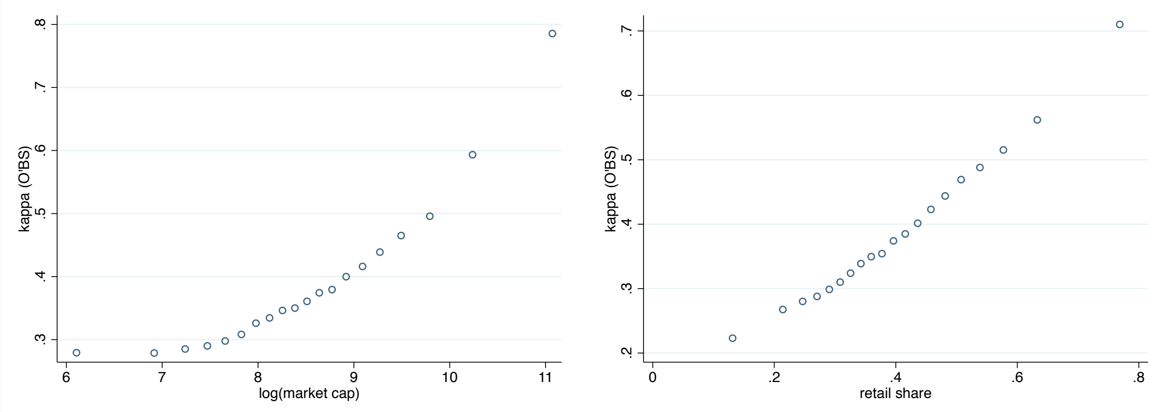


Figure 3: Common Ownership over Time

Figure 4: Heterogeneity in Common Ownership



investment. We find a stark relationship: large market cap firms tend to have substantially higher common ownership weights for other firms. We hypothesize that this is related to retail share. In the common ownership framework, retail investors are infinitesimally small and therefore do not exercise any control over the firm. Large aggregate retail share then tends to inflate the control rights associated with institutional ownership. Indeed, as we see in the second panel of Figure 5, retail share is strongly positively correlated with common ownership weights.

This finding is particularly stark because it relates common ownership to a broader discussion about the separation of control rights and cash-flow rights. In the Berle and Means (1932) tradition "widely held" firms protect shareholders because there is no controlling interest to expropriate them. Porta et al. (1999) contend that this is in fact a rare phenomenon, dominant only in economies like the United States where there are strong protections for minority shareholders. Where such protections are absent, controlling interests may have an incentive to engage in "tunneling," transferring assets towards firms where they have higher cash-flow rights (Johnson et al., 2000). In one respect, the logic of common ownership is similar to tunneling: the ownership structure of the firm leads it to value other firms' profits, perhaps even more than its own. However, the novelty in the common ownership literature is that this needn't take the form of expropriation of minority shareholders. Indeed, in the leading example of price competition in differentiated products, common ownership allows firms in equilibrium to raise prices, which benefits, rather than harms. Instead, consumers are harmed. This has stark implications for that literature: since minority shareholders are, in general, not defrauded, then protection of minority shareholders, who have no incentive to raise lawsuits on common ownership concerns, is an ineffective mechanism in these cases.

Table 3: Correlations with κ

	(1)	(2)	(3)
Year	0.0118*		0.0120*
	(0.0004)		(0.0006)
Market Cap (in logs)	0.0910*	0.0932*	0.0817*
	(0.0033)	(0.0033)	(0.0039)
Retail Share	0.7051*	0.7029*	0.6134*
	(0.0233)	(0.0236)	(0.0268)
Year FE		✓	
Firm FE			✓
SIC Division FE	✓	✓	
N	30796703	30796703	30796703

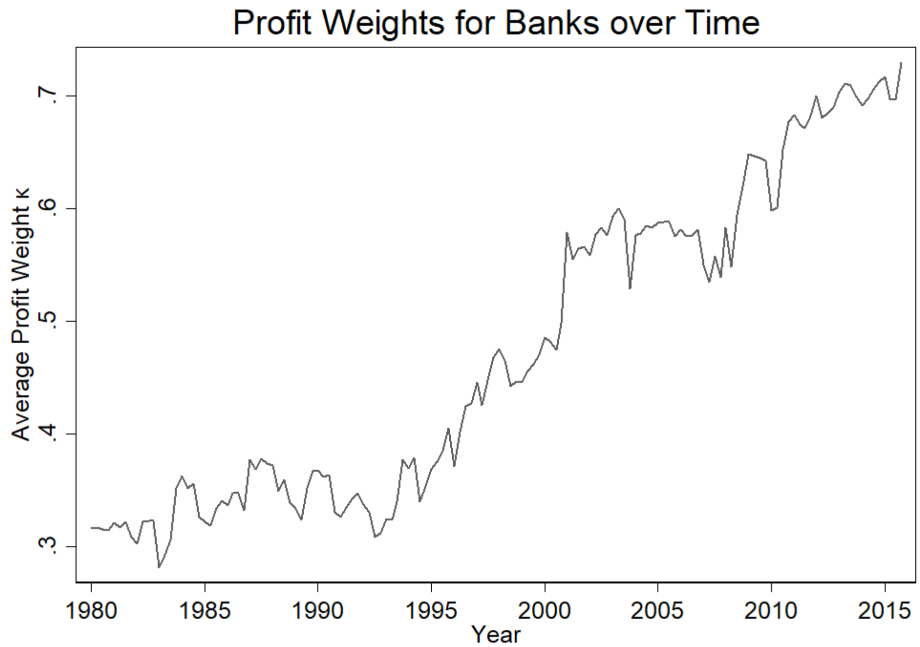
In Table 3 we present these results in regression form. We also add dummy variables for the primary sector of the firm (as identified by COMPUSTAT), which are the most frequent two-digit NAICS codes and the less-frequent together as "other sectors," which is the omitted baseline. All standard errors, noted in parentheses, are clustered at the firm level, and we have also included firm and year fixed effects separately in models (2) and (3). Results are consistent with Figures A-4 and 5: robustly positive with the time trend for OS weights, not for CLWY, but in both cases positively and robustly correlated with market cap and retail share. There is substantially less heterogeneity across sectors – differences are small (less than 0.02) and not strongly significant.

An obvious criticism of the above economy-wide analysis is that a pharmaceutical firm's decisions hardly affect the profits of an airline, so why do these profit weights tell us anything? What are profit weights within relevant product markets? Unfortunately, such an analysis would require market definition, something that we have been able to avoid so far. Below, we present the average profit weight for a specific industry – commercial banks, as defined by SIC code 6020 in Compustat – and show that the qualitative and quantative patterns are similar to those in the S&P 500 as a whole: a large increase over the past few decades.

4.3 Discussion

We believe that the correct object to study under the common ownership hypothesis is not the MHHI or GHHI frequently used in the literature, but instead these κ profit weights implied by ownership patterns. In the rest of the paper we focus on a specific product market – ready-to-eat cereal – but will revisit in that setting the use of MHHI and profit

Figure 5: Profit Weights Among Commercial Banks

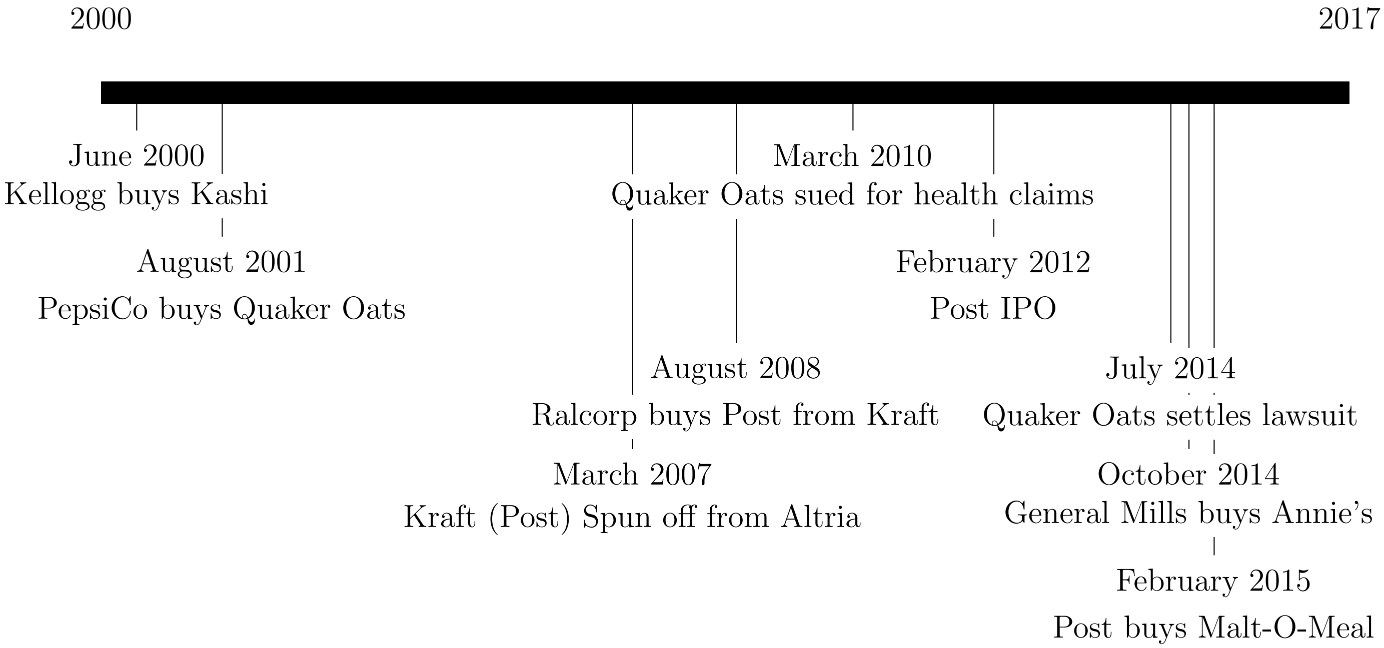


weights in reduced-form analysis, and compare to structural analysis incorporating profit weights directly.

5 Ready-to-Eat Cereal

For the remainder of the paper we consider on the effects of common ownership on pricing behavior, and so we single out a particular market. We focus our empirical exercise on the Ready-To-Eat (RTE) cereal industry for a number of reasons. The first reason is that the industry is highly concentrated with four major players: Kellogg's, General Mills, Quaker Oats, and Post responsible for approximately 85% of the overall marketshare. The second reason is that there are substantial differences in common-ownership across firms. For historic reasons, Kellogg's has large undiversified shareholders while the other firms generally do not. In addition there are a substantial number of transactions both in the ownership space and the product space, particularly involving Post Brands which at various times is a component of the S&P 500 Index, the S&P 400 Midcap Index, and no index at all. The final reason is that there is substantial prior work on the RTE cereal industry which indicates that static

Figure 6: Chronology of RTE: 2000 – 2017



Bertrand-Nash in differentiated products appears to be a reasonable empirical framework.

5.1 The Big Four

Kellogg's: Kellogg's was founded in 1906 by brothers W.K. and J.H. Kellogg, who developed a technique for producing corn flakes that was originally meant for in-house use at the elder brother's Battle Creek Sanitarium. They believed that such a bland and tasteless food, which they called "granula," would reduce impure desires.

Over our sample period the holdings of Kellogg's are stable, however there is one feature that differentiates them from the other three: approximately 20% of Kellogg's shares are held by the W.K. Kellogg Foundation Trust. W.K. Kellogg, the longer-lived brother, established the foundation to ensure that "all children – regardless of race or income – have opportunities to reach their full potential," a marked departure from his elder brother's views, embodied by the Race Betterment Foundation, a now-defunct pillar of the also-defunct eugenics movement. The second largest shareholder is the Gund family which acquired its stake in Kellogg's after selling the decaffeinated coffee brand Sanka to Kellogg's in 1927.

Kellogg's was a member of the S&P 500 over our entire sample and has around 30% marketshare.

Well-known products include Kellogg's Corn Flakes, Fruit Loops, Rice Krispies, Raisin Bran, and Special K. In addition to RTE cereals, Kellogg's also sells other morning foods (e.g., Eggo Waffles, Pop Tarts) and snack foods (e.g., Pringles, Cheez-Its).

General Mills: General Mills' history goes back to the Minneapolis Milling company, founded in 1856. It's main RTE cereal innovation came in 1930, when it developed a technology for "puffing" cereal, as in Kix and Cheerios.

Holdings of General Mills are stable over our sample period, but they did acquire Annie's, a health-conscious brand, in October of 2014, just after our sample concludes. They were included in the S&P 500 for the duration of our sample and has around 30% marketshare.

Well-known products include Cheerios, Chex, Lucky Charms, Total, and Wheaties. Outside of morning foods, they also own Betty Crocker brands, Pillsbury, Nature Valley, Hamburger Helper, Yoplait, and a variety of other food products.

Post: Post also traces its history to the Battle Creek Sanitarium, where founder C.W. Post was a patient. The inaugural cereal, Grape-Nuts, was developed in 1897 under the pretense that it would cure appendicitis. The Postum Cereals company changed its name to the General Foods company in 1929, and was purchased by Phillip Morris (later re-branded as Altria) in 1985 and merged with Kraft in 1989, where it stayed until the beginning of our sample.

Post underwent three major ownership changes during our sample. First, on March 31, 2007, Kraft (which held Post) was spun off from Altria. Next, announced in November 2007, Kraft spun off Post Cereals and the resulting company was sold to Ralcorp Holdings on August 4 of 2008. This transition meant that Post left a S&P 500 company and was now owned by a non-S&P 500 company. Ralcorp Holdings is also a major producer of private label cereals as well as other food products. Finally, announced in July 2011, Ralcorp holdings announced an IPO for the Post Foods Unit, which was successfully spun off on February 7, 2012. The resulting company was not a member of the S&P 500.

Well-known products include Grape-Nuts, Honey Bunches of Oats, and Raisin Bran. In February 2015, Post purchased Malt-O-Meal, a major producer of private label cereals which comprised about 8-9% of the overall market.

Quaker Oats: Quaker Oats is the result of a four-way merger of Midwestern oat mills in 1901. The brand has no affiliation with the Religious Society of Friends (actual Quakers), who regularly display their annoyance at the representation with letter-writing campaigns.

In August of 2001 Quaker oats was purchased by PepsiCo, a S&P 500 company, and it remained in their portfolio for the duration of our sample. From March of 2010 until a settlement in July of 2014, Quaker Oats was subject to a long and public legal battle over the veracity of their health claims. They did not claim that their products cured appendicitis or moral impurity; none the less, the legal battle may have contributed to a decline in sales.

Well-known brands include Cap'n Crunch, and Life and Quaker Oats has around 8-9% of the cereal market. The company also produces other morning foods (Oatmeals, and Aunt Jemima branded foods) as well as other food products, but should be considered in the larger PepsiCo setting, where it makes up a relatively small fraction of sales (2-3%).

5.2 Data Sources: Ownership Data

Our ownership data is a subset of the data described in Section 4.1. We are interested in the following firms, which at some point in 2004-2017 offered products in the ready-to-eat cereal market: General Mills (GIS), Kellogg (K), Kraft (KRFT: Q1 2013 - Q2 2015), Mondelez (MDLZ), Altria Group (MO), PepsiCo (PEP), Philip Morris (PM: Q4 2008 - Q2 2017), Post Holdings (Q1 2012 - Q2 2017), and Ralcorp (RAH: Q1 2008 - Q4 2011). We made a small number of corrections to the Capital IQ data to address potential double counting of large private holdings as described in Data Appendix B.

Table 4 provides summary statistics on the common ownership data. There are some important patterns to point out. The first is that Vanguard appears to be increasing its holdings across all firms over time. In part this is driven by the growing share of Vanguard within the index fund market. The second is that between 2004 and 2010 there is a reallocation from Barclays Global Investors (BGI) and Blackrock which acquired the BGI exchange-traded-fund (ETF) business in June of 2009. This made Blackrock the largest player in the ETF market.[8] State Street is another large player in the ETF market and also sees their ownership stakes increasing over time. Another large player is FMR LLC, which is better known as the financial entity behind Fidelity which is a major player in both actively managed and index funds. Capital Research, the parent of company of the American Funds family (primarily actively managed) is another major player particularly in the early periods. We provide a more detailed accounting of ownership stakes over time by major investors in the Appendix.

Table 4: Top 5 Owners of Major Firms, 2004-2016

General Mills (GIS)

2004		2010		2016	
Capital Research and Management	7.28%	BlackRock, Inc	8.70%	BlackRock, Inc	7.36%
Barclays Global Investors	3.24%	State Street Global Advisors	5.92%	The Vanguard Group	6.92%
Wellington Management Group	3.06%	The Vanguard Group	3.56%	State Street Global Advisors	6.14%
State Street Global Advisors	2.48%	MFS	2.65%	MFS	3.37%
The Vanguard Group	1.95%	Capital Research and Management	2.43%	Capital Research and Management	2.12%

Kellogg (K)

2004		2010		2016	
W.K. Kellogg Foundation	29.87%	W.K. Kellogg Foundation	22.94%	W.K. Kellogg Foundation	19.75%
Gund Family	7.26%	Gund Family	8.65%	Gund Family	7.68%
Capital Research and Management	2.83%	Capital Research and Management	3.54%	The Vanguard Group	4.97%
Barclays Global Investors	2.81%	BlackRock, Inc	2.97%	BlackRock, Inc	4.64%
W.P. Stewart & Co.	2.63%	The Vanguard Group	2.42%	MFS	3.51%

Quaker Oats, a Unit of PepsiCo (PEP)

2004		2010		2016	
Barclays Global Investors	4.40%	BlackRock, Inc	4.64%	The Vanguard Group	6.72%
State Street Global Advisors	2.81%	Capital Research and Management	4.37%	BlackRock, Inc	5.63%
FMR LLC	2.74%	The Vanguard Group	3.64%	State Street Global Advisors	3.98%
The Vanguard Group	2.08%	State Street Global Advisors	3.19%	Wellington Management Group	1.48%
Capital Research and Management	1.82%	Bank of America	1.63%	Northern Trust	1.37%

Post Brands, a Unit of Altria (2004, MO), Ralcorp (2010, RAH), and Post Holdings (2016, POST)

2004		2010		2016	
Capital Research and Management	7.37%	FMR LLC	10.18%	Wellington Management Group	9.63%
State Street Global Advisors	3.61%	BlackRock, Inc	8.35%	BlackRock, Inc	8.42%
Barclays Global Investors	3.51%	The Vanguard Group	3.57%	FMR LLC	7.24%
FMR LLC	2.60%	Baron Capital Group	3.39%	The Vanguard Group	6.93%
AllianceBernstein L.P.	2.25%	Steinberg Asset Management	2.68%	Tourbillon Capital Partners	6.89%

5.3 Data Sources: Sales and Product Data

Our primary data source for sales and prices of ready-to-eat (RTE) cereal comes from the Kilts Nielsen Scanner Dataset. The data are organized by store, week, and UPC code. For each store-week-upc we observe unit sales as well as a measure of "average price" which is revenue divided by sales. Prices may vary within a UPC-store-week for a number of

[8] Azar et al. (2017) use this event as an instrument for changes in ownership as it substantially increases the holdings of Blackrock.

reasons: the first is that price changes may occur within the middle of the reporting week, the second is that some consumers may use coupons; according to the documentation [CITE] retailer coupons or loyalty card discounts are included in the "average price" calculation while manufacturer coupons are not. Because there is little price variation across stores within the same chain (DellaVigna and Gentzkow, 2017), we aggregate unit sales and revenues to the DMA-chain level.[9]

While the Nielsen Scanner dataset collects data from stores across the entire United States, we focus on six designated market areas (DMAs) where the estimated coverage of Nielsen reporting supermarkets is high. We focus exclusively on conventional supermarket sales (F) stores, and exclude pharmacies which sometimes sell RTE cereal (D stores) and mass-market (M stores). In Table 5 we report the number of supermarkets in our dataset for the six DMAs we analyze. We chose these markets because these are markets where Nielsen reporting stores comprise a large overall share of all supermarkets within the DMA. This has the advantage that we gain a relatively complete picture of prices and quantities within the DMA, though it has the disadvantage that it tends to select DMAs with dominant chain retailers (who report to Nielsen). The Nielsen Scanner Dataset contains data from 2006-2016. We exclude 2006, and focus on data from 2007-2016 only because the set of stores observed in 2006 is differs from the set of stores observed in subsequent years. This leaves us with 3440 UPCs, in 41 chains, over 522 weeks for a total of over 4.8 million observations, and around 200-300 products per DMA-chain-week.

DMA	# of stores	# of chains	% Coverage
Redacted			

Table 5: Number of Stores and Market Coverage by DMA for (F)ood Stores
% Coverage reports share of Nielsen's calculated All Commodity Volume (ACV) by DMA-Channel.

One drawback of the Nielsen Scanner dataset is that it provides only limited product-level information. At the UPC level we don't observe much beyond the brand name (e.g. Honey Nut Cheerios), the manufacturer names (General Mills), and the package size (14 oz. box). To overcome this limitation we collect nutritional information from the Nutritionix Database. This database is organized by UPC code and was designed to provide API access for various fitness tracking mobile apps. It encodes the nutritional label on the product packaging

[9] We provide supporting evidence in favor of chain-level pricing, including the heatmap plots in the Appendix XXX.

(serving size, calories, sugar, fat, vitamin content, ingredient lists). We merged the Nielsen UPC information with the nutritional label information from Nutritionix. A large share of products in our dataset (9-10% by volume) are private-label brands. For these products, we do not have UPC codes which we can match to the Nutritionix database.[10] Instead, we must match these products to the most similar branded product *HNY TSTD O'S* to *Honey-Nut Cheerios* and use the nutritional information from the branded product. For some private label products we cannot identify the most similar brand (e.g. *CTL BR C-M-C RTE*), rather than dropping these products we impute product characteristics using averages.

Throughout the paper our preferred definition of quantity will be in "servings". The serving size is generally measured by weight in grams or ounces and for each box of cereal we convert the package size (in ounces) to the number of equivalent servings. We convert all products to single-serving equivalents (in terms of nutritional information, price and quantity). This has advantages and disadvantages. Nutrient dense cereals generally display smaller serving sizes (by weight) which lead to much smaller serving sizes by volume. There is some evidence that serving sizes are chosen so that caloric content falls in the (100-150) range rather than measuring typical serving sizes by consumers.[11] Another issue is that package sizes are declining over time. From the beginning of our sample in January 2007 until the end of our sample in December 2015, the typical box of cereal shrank by approximately 8% going from around 13.3 ounces to 12.2 ounces servings.[12] We illustrate this in Figure 7.

Because there are a large number of product characteristics (19), and we aren't interested in nutritional aspects of products *per se* we consolidate the nutritional information into a number of *factors*. The idea is to reduce the dimension of the characteristic space while preserving the variation across products. This lets us measure which products are more (or less) similar based on nutritional content. We elaborate on this process in Appendix A.1.

Redacted

Figure 7: Average Package Size Over Time

[10] We believe these UPCs are obfuscated by Nielsen to prevent researchers from de-identifying chains. The true UPC code could identify the product as *Chain X Brand Honey Toasted O's*.

[11] [CITE:CR] conducted a survey and found that 92% of consumers exceed the posted serving size when pouring bowls of cereal. The average overpour on Cheerios was 30%-130% while it was even greater for denser cereals like Museli or Granola where the average overpour was 282%.

[12] The fact that weight declines more quickly than the number of servings provides some evidence that over time consumers substitute away from dense cereals such as Museli and Granola and towards lighter cereals such as Flakes or O's.

6 Common Ownership and Pricing

Much attention in the common ownership literature has been paid to the Modified Herfindahl-Hirschman Index (MHHI) concentration measure, which is derived from a Cournot oligoopoly model of competition in O'Brien and Salop (2000).[13] MHHI extends the traditional concept of HHI to incorporate common ownership, and is defined as follows:

$$\max_{q_f} \pi_f(q_f, q_{-f}) + \sum_g \kappa_{fg} \pi_g(q_f, q_{-f})$$

After taking FOC we get:

$$\frac{P_f - MC_f}{P_f} = \frac{1}{\eta} \sum_g \kappa_{fg} s_g$$

Which gives the share weighted average markup of:

$$\begin{aligned} \sum_f s_f \frac{P_f - MC_f}{P_f} &= \frac{1}{\eta} \underbrace{\sum_f \sum_g \kappa_{fg} s_g s_f}_{MHHI} \\ \text{where } MHHI &= \underbrace{\sum_f s_f^2}_{HHI} + \underbrace{\sum_f \sum_{g \neq f} \kappa_{fg} s_f s_g}_{\Delta MHHI} \end{aligned} \tag{3}$$

The Price Pressure Index (PPI) is similarly defined for differentiated Bertrand competition. We consider the objective function for firm f when setting the price p_j holding fixed the prices of all other products p_{-j}. As firm f raises the price p_j some consumers substitute to other brands owned by f: $k \in \mathcal{J}_f$ on which it receives full revenue, and substitute brands owned by competing firms g: $k' \in \mathcal{J}_g$ for which it acts as if it receives a fraction of the revenue κ_{fg}:

$$p_j q_j(p_j, p_{-j}) - c_j(q_j) + \sum_{k \in \mathcal{J}_f} p_k q_k(p_j, p_{-j}) - c_k(q_k) + \sum_g \kappa_{fg} \cdot \left(\sum_{k' \in \mathcal{J}_g} p_{k'} q_{k'}(p_j, p_{-j}) - c_{k'}(q_{k'}) \right)$$

[13]Originally the MHHI was derived by Bresnahan and Salop (1986) in the context of a joint-venture.

After taking FOC we get:

$$q_j + (p_j - mc_j)\frac{\partial q_j}{\partial p_j} + \sum_{k \in \mathcal{J}_f} (p_k - mc_k)\frac{\partial q_k}{\partial p_j} + \sum_g \kappa_{fg} \cdot \left(\sum_{k' \in \mathcal{J}_g} (p_{k'} - mc_{k'})\frac{\partial q_{k'}}{\partial p_j} \right) = 0 \qquad (4)$$

Now it is helpful to do two things: (1) divide through by $-\frac{\partial q_j}{\partial p_j}$; (2) define the diversion ratio $D_{jk} = -\frac{\frac{\partial q_k}{\partial p_j}}{\frac{\partial q_j}{\partial p_j}}$. We can then solve for p_j (this expression is known as the PPI):

$$\begin{aligned} p_j &= -q_j / \frac{\partial q_j}{\partial p_j} + mc_j + \sum_{k \in \mathcal{J}_f} (p_k - mc_k) D_{jk} + \sum_g \kappa_{fg} \cdot \left(\sum_{k' \in \mathcal{J}_g} (p_{k'} - mc_{k'}) D_{jk'} \right) = 0 \\ p_j &= \frac{\epsilon_j}{\epsilon_j - 1} \left[mc_j + \sum_{k \in \mathcal{J}_f} (p_k - mc_k) D_{jk} + \sum_g \kappa_{fg} \cdot \left(\sum_{k' \in \mathcal{J}_g} (p_{k'} - mc_{k'}) D_{jk'} \right) \right] \qquad (5) \end{aligned}$$

The result is that the price p_j has the usual elasticity rule applied to marginal cost, but marginal cost has been augmented in two ways: (1) to include the profit margin and diversion ratio to brands owned by the same firm, and (2) to include κ_{fg} weighted profit margins and diversion ratios for products with common owners. It is important to note that we can only solve for p_j under the assumption that p_{-j} remains fixed and does not respond, thus this does not necessarily represent an equilibrium price. What is more robust to equilibrium price responses is the notion that much like the difference between single-product and multi-product monopoly pricing, common ownership raises the opportunity cost of selling product j because some fraction of sales are "re-captured" to commonly owned products.

6.1 Analytic Example

In order to highlight some of the challenges associated with using cross-sectional variation to identify the relationship between common-ownership and prices we present an extremely simple analytic example. We begin with two firms playing a static differentiated Bertrand-Nash (in prices) game facing linear demand curves given by:

$$\begin{aligned} q_1 &= a_1 - \alpha_1 \cdot p_1 + \beta_1 \cdot p_2 \\ q_2 &= a_2 - \alpha_2 \cdot p_2 + \beta_2 \cdot p_1 \end{aligned}$$

To further simplify we assume symmetric marginal costs of zero so that $mc_1 = mc_2 = 0$. This allows us to write best-responses assuming that each firm maximizes their own profits and ignores the profits of the rival:

$$
\begin{aligned}
p_1(p_2) &= \frac{a_1 + \beta_1 \cdot p_2}{2\alpha_1} \\
p_2(p_1) &= \frac{a_2 + \beta_2 \cdot p_1}{2\alpha_2}
\end{aligned}
$$

Which can easily be solved for a closed form solution:

$$
p_i = \frac{2a_i\alpha_j + \beta_i a_j}{4\alpha_i\alpha_j - \beta_i\beta_j}
$$

It is helpful to point out that p_i is increasing in both intercepts (a_i, a_j) so that increasing either demand intercept monotonically raises the prices of both goods.

Now we conduct our cross-market exercise. We set $(a_2, \alpha_1, \alpha_2, \beta_1, \beta_2) = (100, 1, 1, 0.5, 0.5)$ and parametrize each "market" by a different demand intercept a_1 which we vary over $[80, 100]$ and so that:

$$
\begin{aligned}
q_1 &= a_1 - p_1 + \frac{1}{2} \cdot p_2 \quad \text{for } a_1 \in [80, 100] \\
q_2 &= 100 - p_2 + \frac{1}{2} \cdot p_1
\end{aligned}
$$

We can calculate $p_1^*(a_1), p_2^*(a_1)$ as well as $HHI(a_1)$. We also assume that we (incorrectly) believe that $\kappa_{12} = \kappa_{21} = 0.5$ for all markets and calculate $\Delta MHHI(a_1) = 2\kappa \cdot s_1(a_1)\, s_2(a_1)$. In fact, the equilibrium prices and quantities are all calculated under the assumption that $\kappa = 0$ or that the firm ignores the rivals profits. In Figure 8, we plot both prices and $\Delta MHHI$ as a function of the demand intercept a_1 holding all other quantities fixed. For $a_1 < 100$ we find that there is a positive correlation between prices (p_1 in blue and p_2 in red) and $\Delta MHHI$, while for $a_1 > 100$ we find that there is a negative correlation between prices and $\Delta MHHI$. This illustrates a number of important points when considering the empirical relationship between price and $\Delta MHHI$: (1) the relationship may be non-monotonic; (2) differences in demand across markets may lead to spurious correlations in either direction; (3) cross-market variation will come largely through demand conditions and marketshares as κ is set nationally and varies only in the time series.

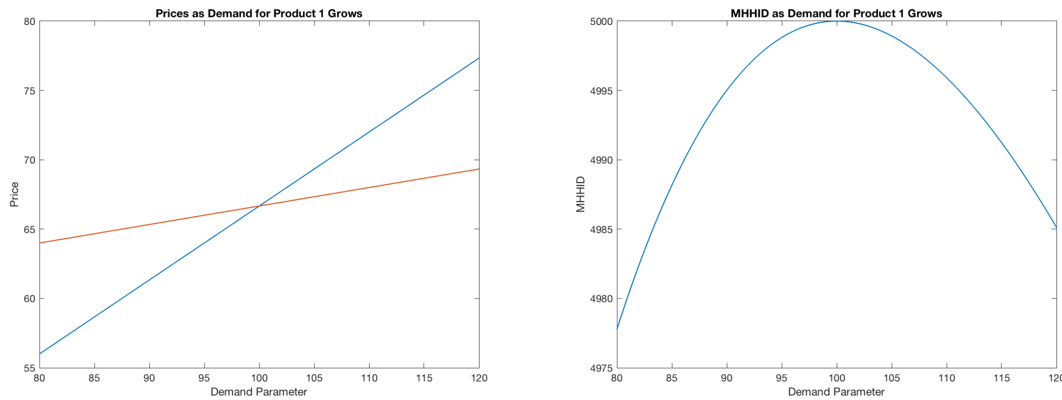


Figure 8: Analytic/Numerical Example: Varying Demand Intercept

6.2 Identification Discussion

The existing empirical literature on common ownership generally relies on structure-conduct-performance (SCP) style regressions of the form:

$$\log p_{jmt} = \gamma_{mj} + \beta_1 HHI_{mt} + \beta_2 \Delta MHHI_{mt} + \beta_3 s_{jmt} + c_1 t + c_2 t^2 + \ldots + \epsilon_{jt} \tag{6}$$

where j subscripts a firm (or a product), m subscripts a market, and t subscripts a period. In this regression: (1) $\beta_1 > 0$ (concentration is associated with higher prices); (2) $\beta_2 > 0$ (common ownership is associated with higher prices); (3) $\beta_3 < 0$ (larger share is associated with lower prices). The last point is generally either believed to be evidence that profitability is associated with efficiency (Demsetz, 1973) or merely evidence that demand slopes downwards. Azar et al. (2017) include various controls in their study of airlines such as route specific fixed effects as well as carrier specific fixed effects (though they exclude s_{jmt}) when they run a regression in the form of (6).

There is a long literature criticizing regression equations of the form in (6). One major critique is that outside of some very restrictive assumptions about symmetric homogenous Cournot competition there may not be a theoretical relationship between prices or profitability and HHI or $\Delta MHHI$. Additionally because both HHI and $\Delta MHHI$ are functions of marketshares, there is an endogeneity problem. Particularly when the left-hand-side variable is price, this becomes a regression of prices on some functions of quantities which may represent either a supply curve or a demand curve. Furthermore any strategy must instrument for marketshare in both HHI and $MHHI$, and should also instrument for κ as ownership is also likely to be endogenous.

As pointed out by O'Brien (2017) we can actually try to generate a true "reduced form" equation in the common ownership framework:

$$
\begin{aligned}
y_f &= f_f(y_{-f}, X, \kappa_f) && \text{for } f = 1,\ldots,F && \text{structural equations} \\
y_f &= g_f(X, \kappa) && \text{for } f = 1,\ldots,F && \text{reduced form equations} \\
y_f &= h_f\left(X, \sum_g s_g^2, \sum_f \sum_{f \neq g} \kappa_{fg} s_f s_g\right) && \text{for } f = 1,\ldots,F && \text{SCP equations}
\end{aligned}
\tag{7}
$$

To go from the structural equation to the "reduced form" we need to be able to solve the system of equations for policies from best-responses. This generally requires some assumptions to guarantee that such. The "true" reduced form should depend on the entire matrix of κ's. It is also important to note that the SCP regressions are neither a reduced form nor a structural equation, instead they rely on particular functions and interactions of (s_f, s_g, κ_{fg}).

We also observe that if one relies on panel data in order to estimate κ_{fg} there are some potential problems. The first is that κ_{fg} will vary only over time t but not across markets m. We expect κ_{fg} to vary with holdings of institutional investors from quarter to quarter, but there is no mechanism which allows firm f to place a different weight on the profits of firm g in one market than it does in another. This is problematic for many identification strategies. We cannot rely on cross market variation (or cross market instruments) in order to identify κ. Moreover, we are not really interested in the effect of κ interacted with marketshare across markets, but rather the effect of κ on prices directly (the "true" reduced form). Our analytic/numerical example has already demonstrated why cross-market identification strategies may be dangerous or misleading in this context.

7 Descriptive Evidence for Common Ownership

7.1 κ weights, Price Indices, and Concentration Measures

In Figure 9 we plot the κ weights that each firm places on his competitor's profits.

For example, the top left pane shows the weight that Kellogg's puts on the profits of their competitors. Notice that the weight Kellogg's puts on its own profit is normalized to one and constant over time. The weights are similar across competitors and slowly growing over time from around 8% to 20%. These relatively small weights are due to the large undiversified Kellogg's shareholders (Kellogg Family Foundation and Gund Family). Contrast this with General Mills in the second pane. General Mills places between 60-80% weight on the profits of Quaker Oats and Post as it does on its own profits, with substantial variation across time. It places slightly less weight on the profits of Kellogg's because of less overlapping ownership, though still more weight 40-60% than Kellogg's places on the profits of General Mills. Quaker Oats (a division of PepsiCo) occasionally places more weight $\kappa > 1$ on competitor's (General Mills and Post) profits than it does on its own profits. Quaker Oats puts somewhat less weight (though still $\kappa > 0.6$ on the profits of Kellogg's which has less overlap in ownership. Post generally puts less weight on each of its competitor's profits over time as Post is transitions from an S&P 100/500 component, to an S&P 400 Midcap Index Component, and briefly after its 2012 IPO is not included in any index, before rejoining the S&P 400 Midcap Index.

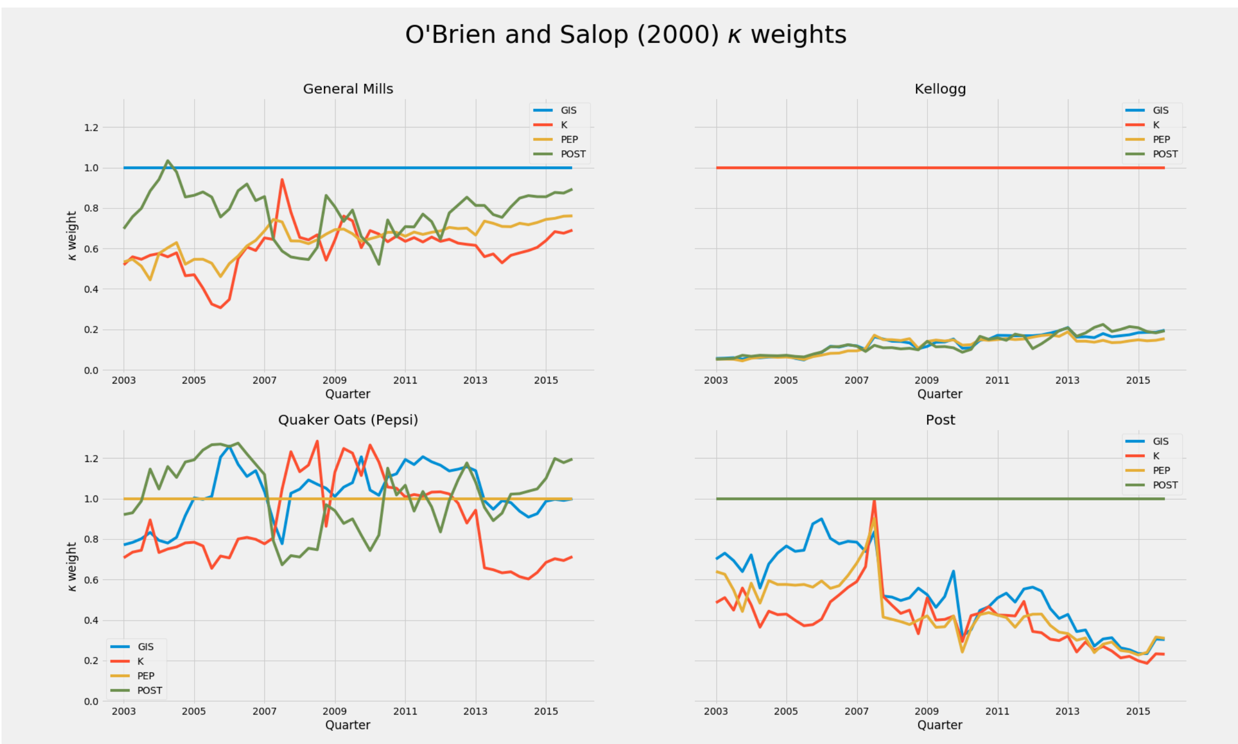


Figure 9: κ Profit Weights for Ready-To-Eat Cereal

The next thing we calculate are price indices over time. We construct these price indices by defining a fixed basket of goods based on the top 250 products at the UPC level and tracking those products over time. Before computing any price indices, we smooth our data out to be quarterly so that the "price" is in fact total quarterly revenue divided by quarterly servings at the DMA level. We construct our price indices as follows:

$$P_I(t) = \frac{\sum_{j \in T_0} p_{jt} \cdot q_{j0}}{\sum_{j \in T_0} q_{j0}}, \quad P_V(t) = \frac{\sum_{j \in T} p_{jt} \cdot q_{jt}}{\sum_{j \in T_0} q_{jt}}$$

where T_0 is the set of the 250 best selling products in period 0 and q_{j0} are the sales of product j in period $t = 0$ as measured by number of servings. Thus $P_I(t)$ represents the weighted average serving price for the set of products purchased in period $t = 0$. We consider three ($t = 0$) periods for measuring our sales: (1) 2008 Quarter 3, (2) 2010 Quarter 1, and (3) 2015 Q3. We chose these to be towards the beginning, middle and end of our sample. We also report weighted average prices for a variable basket of products which we label $P_V(t)$. All of these measures have challenges in that not all UPCs are available during all periods. The primary source of this missing data is that product packaging changes and products are assigned new UPCs. For example, the package size "shrinks" from 14oz to 12oz.

We plot these price indices at the DMA level in Figure 10. Each DMA has its own set of q_{0t} weights as the bundle of products purchased in each city may vary. Price per serving varies from between 20-27 cents per serving across cities and over time. With the exception of Boston, most cities experience a substantial dip in prices in early 2010 before eventually recovering. There is substantial cross-sectional variation in prices with Boston and Charlotte seeing above average prices and Denver and Phoenix seeing below average prices. We cannot say how much of the cross-sectional variation comes from different regional chains rather than regional preferences. We also report the same set of index prices for the entire sample (rather than broken out by DMA) in Figure 11. This figure shows more clearly that there is a precipitous drop in prices in early 2010 and before recovering at the end of 2011 and leveling off afterwards. This recovery in prices appears to begin around the same time as the Blackrock acquisition of Barclay's Global Investors ETF business, which we denoted with a vertical green line. If this event is both the source of a substantial increase in common ownership and the cause of the price recovery in 2010-2011, then this would be evidence in favor of the common ownership hypothesis. We also plot black dashed lines for each of the major events in Post's timeline (sale to Ralcorp from Kraft, IPO, and acquisition of Malt-O-Meal). We also denote the General Mills acquisition of Annie's Homegrown (the largest independent producer of premium, organic RTE cereal) with the black dotted line.

Redacted

Figure 10: Price Indices Across Time and Markets

Redacted

Figure 11: Price Indices for All DMAs

We document some additional facts around the substantial price decline in early 2010 and subsequent recovery. We plot the marketshare of private label products across all DMAs and then later broken out by DMA in Figure 12. One narrative that has some support in popular accounts [CITE] is that during the Great Recession there was an increase in the private label share (from around 15 to 18.5% of the market), in early 2010 the four dominant manufacturers responded with substantial price cuts causing the private label share to fall, before abandoning the price cuts in 2011 whereby price indices and private label share recovered to pre "Price War" levels before private label share declined as the economy continued to improve.

Redacted

Figure 12: Private Label Marketshare

We can also use marketshares to construct quarterly concentration measures (such as the HHI). We present plots of HHI across time and markets in Figure 13. We include the same set of vertical lines to denote the same set of transactions and events as before. Because we don't necessarily know which manufacturers produce the private label products (around 50% of the private label market is produced by Malt-O-Meal but we don't know which products), we instead assume that each privately label product is produced by a different manufacturer. As we see in Figure 13 there is substantial cross market concentration in HHI with Denver and Phoenix being relatively unconcentrated (HHI $<$ 1800) while Chicago is highly concentrated (HHI $>$ 2500). The concentration more less mirrors the inverse of the private label share (Chicago and Charlotte are more concentrated and have a lower private label share). When we look at HHI averaged across all markets we see relatively little response to the Blackrock/BGI event (as we would expect), we see a substantial increase in HHI after the Post/Malt-O-Meal and General Mills/Annie's Homegrown acquisitions towards the end of the sample. We also see a substantial decline in HHI around the same time as Kraft sold Post to Ralcorp. Across time we rarely see more than a 150 point change in the national aggregate HHI.

Redacted

Figure 13: Herfindahl-Hirschman Index (HHI) over time

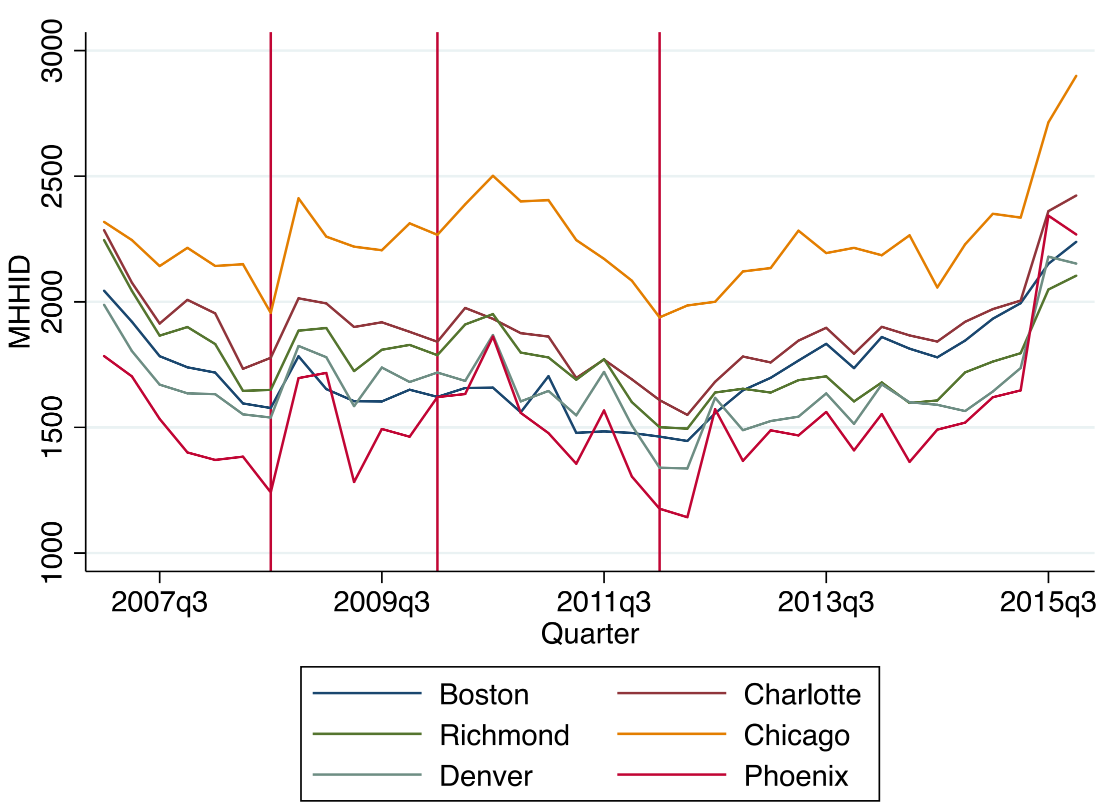


Figure 14: $\Delta MHHI$ over time

The main point of Figure 13 is to demonstrate that other than the spike in private label sales during 2009, there isn't much variation within a DMA over time, but there is much larger variation across DMAs in market concentration. This cross market variation in HHI is likely to drown any time series variation in κ when we construct $\Delta MHHI$. Indeed when we construct the $\Delta MHHI$ and plot it across DMAs in Figure 14 we find the majority of the variation is across market rather than within market over time. We find that the $\Delta MHHI$ is approximately the same magnitude as the HHI: around 1500 for Phoenix and around 2400 for Chicago.

7.2 Regression Based Evidence

We now perform the same set of Price-Concentration regressions as the previous literature: Azar et al. (2017), Azar et al. (2016), Kennedy et al. (2017), Gramlich and Grundl (2017). We do not call these a "reduced form" as they are not necessarily the reduced form of any particular equilibrium model. Recall equation (6):

$$\log p_{jmt} = \gamma_{mj} + \beta_1 HHI_{mt} + \beta_2 \Delta MHHI_{mt} + \beta_3 s_{jmt} + c_1 t + c_2 t^2 + \ldots + \epsilon_{jt}$$

We report our regression results in Tables 6 - 9. We perform these regressions at two levels of aggregation. The first is at the manufacturer-chain-DMA-quarter level, and the second is at the manufacturer-DMA-quarter level. While we might want to use higher frequency time series data, because we only observe 13-F filings once per quarter such variation would be dubious to use to identify common ownership effects. We use both prices and log prices as our left hand side variable. When we construct prices we use the $P_V(t)$ version of the price index (just total revenue divided by total servings). We include fixed effects at the dma-level, chain-dma-level, manufacturer level, and a quadratic time trend. In some specifications we also include a cubic polynomial in HHI, in case $\Delta MHHI$ is picking up nonlinearities in the price-concentration relationship. In all of our specifications there is a negative and significant relationship between the $\Delta MHHI$ and price. This suggests that additional overlapping ownership is correlated with *lower prices* rather than *higher prices*. We don't interpret this as causal mechanism. That is, it is inappropriate to conclude that common ownership leads to *lower prices*. Instead we offer our numerical example from Section 6.1 as an explanation: cross sectional variation in demand provides variation in marketshares. When we interact these marketshares with κ_{fg} we can produce spurious positive or negative correlations between $\Delta MHHI$ and price.

Table 6: DMA Level MHHI Regressions p

	(1)	(2)	(3)	(4)	(5)
	Price	Price	Price	Price	Price
	b/se	b/se	b/se	b/se	b/se
hhi_servings	0.0123***	0.0221***	-0.0062***	0.0021	0.0028*
	(0.0029)	(0.0031)	(0.0019)	(0.0015)	(0.0016)
mhhid_servings	-0.0261***	-0.0254***	-0.0351***	-0.0232***	-0.0243***
	(0.0034)	(0.0034)	(0.0020)	(0.0016)	(0.0016)
DMA FEs	No	Yes	No	Yes	Yes
Manufacturer FEs	No	No	Yes	Yes	Yes
Quadratic Time Trend	Yes	Yes	Yes	Yes	Yes
Cubic HHI	No	No	No	No	Yes
R^2	0.110	0.204	0.831	0.903	0.904
N	1294	1294	1294	1294	1294

* p¡0.10, ** p¡0.05, *** p¡0.01

Table 7: DMA Level MHHI Regressions $\log p$

	(1)	(2)	(3)	(4)	(5)
	Log(Price)	Log(Price)	Log(Price)	Log(Price)	Log(Price)
	b/se	b/se	b/se	b/se	b/se
hhi_servings	0.0396***	0.0730***	-0.0141*	0.0121*	0.0028*
	(0.0115)	(0.0123)	(0.0074)	(0.0062)	(0.0016)
mhhid_servings	-0.1092***	-0.1059***	-0.1388***	-0.0959***	-0.0243***
	(0.0134)	(0.0136)	(0.0075)	(0.0064)	(0.0016)
DMA FEs	No	Yes	No	Yes	Yes
Manufacturer FEs	No	No	Yes	Yes	Yes
Quadratic Time Trend	Yes	Yes	Yes	Yes	Yes
Cubic HHI	No	No	No	No	Yes
R^2	0.124	0.193	0.846	0.899	0.904
N	1294	1294	1294	1294	1294

* p¡0.10, ** p¡0.05, *** p¡0.01

d

As a final exercise we run a "true reduced form" regression as suggested by O'Brien (2017) of the form in equation (7). Here an observation is a manufacturer-chain-dma-quarter. The variable of interest is κ. Theoretically it should be the entire matrix K which we summarize by using the arithmetic or geometric mean of the off-diagonal κ entries. We report the results of the regression of average prices on these functions of K in Table 10. Once again we observe a negative (and significant) relationship between the common ownership profit weights κ and the prices p_{jmt}. For some specifications the results are barely (in)significant.

Table 8: DMA-Chain Level MHHI Regressions p

	(1) Price b/se	(2) Price b/se	(3) Price b/se	(4) Price b/se	(5) Price b/se	(6) Price b/se
hhi_servings	0.0130***	0.0234***	-0.0067***	0.0021	0.0110***	0.0105***
	(0.0017)	(0.0018)	(0.0016)	(0.0015)	(0.0013)	(0.0014)
mhhid_servings	-0.0269***	-0.0265***	-0.0358***	-0.0240***	-0.0181***	-0.0192***
	(0.0019)	(0.0019)	(0.0016)	(0.0015)	(0.0013)	(0.0014)
share_servings					-0.0026***	-0.0026***
					(0.0001)	(0.0001)
DMA FEs	No	Yes	No	Yes	Yes	Yes
Retailer FEs	No	No	Yes	Yes	Yes	Yes
Quadratic Time Trend	Yes	Yes	Yes	Yes	Yes	Yes
Parent FEs	Yes	Yes	Yes	Yes	Yes	Yes
Cubic HHI	No	No	No	No	No	Yes
r2_a	0.229	0.295	0.720	0.767	0.820	0.820
N	5173	5173	5173	5173	5173	5173

* p¡0.10, ** p¡0.05, *** p¡0.01

Table 9: DMA-Chain Level MHHI Regressions $\log p$

	(1) Log(Price) b/se	(2) Log(Price) b/se	(3) Log(Price) b/se	(4) Log(Price) b/se	(5) Log(Price) b/se	(6) Log(Price) b/se
hhi_servings	0.0495***	0.0857***	-0.0105	0.0181***	0.0578***	0.0559***
	(0.0068)	(0.0074)	(0.0066)	(0.0064)	(0.0057)	(0.0060)
mhhid_servings	-0.1216***	-0.1204***	-0.1523***	-0.1097***	-0.0833***	-0.0892***
	(0.0079)	(0.0081)	(0.0065)	(0.0064)	(0.0056)	(0.0059)
share_servings					-0.0116***	-0.0117***
					(0.0003)	(0.0003)
DMA FEs	No	Yes	No	Yes	Yes	Yes
Retailer FEs	No	No	Yes	Yes	Yes	Yes
Quadratic Time Trend	Yes	Yes	Yes	Yes	Yes	Yes
Parent FEs	Yes	Yes	Yes	Yes	Yes	Yes
Cubic HHI	No	No	No	No	No	Yes
r2_a	0.267	0.312	0.722	0.754	0.814	0.814
N	5173	5173	5173	5173	5173	5173

* p¡0.10, ** p¡0.05, *** p¡0.01

Table 10: $\log p$ on κ regression

	Mean	Geometric Mean
	(1)	(2)
Market Share	-0.0026*	-0.0026*
	(0.0001)	(0.0001)
κ	-0.0771*	
	(0.0115)	
κ		-0.0951*
		(0.0120)
Retailer, Parent, DMA FE	✓	✓
Quadratic Time Trend	✓	✓
N	5173	5173

8 Structural Model of Cereal Demand and Supply

8.1 Demand

We define the utility of consumer i for product j and store-week t as:

$$u_{ijt} = \delta_{jt} + \mu_{ijt} + \varepsilon_{ijt}$$

As per usual we assume that consumers also possess an outside option $j = 0$ which provides utility $u_{i0t} = \varepsilon_{i0t}$. We assume ε_{ijt} follows a Type I extreme value (Gumbel) distribution so that the predicted purchase probabilities can be written:

$$s_{ijt}(\delta_t, \mu_i) = \frac{e^{\delta_{jt}+\mu_{ijt}}}{1 + \sum_k e^{\delta_{kt}+\mu_{ikt}}} \tag{8}$$

$$s_{jt}(\delta_t, \theta) = \int s_{ijt}(\delta_t, \mu_i) f(\mu_{ijt}|\theta) \partial\mu_i \tag{9}$$

and match overall sales:

$$q_{jt} = M_t \cdot s_{jt}(\delta_t, \theta) \quad \text{or} \quad \frac{q_{jt}}{M_t} = s_{jt}(\delta_t, \theta)$$

We define a product j as a brand-size combination (e.g. a 14 oz. box of Honey-Nut Cheerios), and a market t as a DMA-chain-week. We aggregate sales across all stores within the same retail chain and DMA (ie: all *Chain X's in Chicago*). We measure q_{jt} or s_{jt} in terms of serving equivalents of cereal as in (Nevo, 2001). Likewise we consider the per-serving price as p_{jt}. We define the market somewhat differently than in previous work examining supermarket purchases.[14] Our definition for marketsize, M_t,is proportional to the number of consumers who walk into the store each week which we measure by tracking weekly store-level purchases of two of the most frequently purchased consumer staples (milk and eggs).[15] We provide a detailed description for how we construct M_t in the Appendix.

We maintain the fiction that each "individual" arrives at the store and chooses to purchase a single serving of cereal at the per-serving price, or selects the outside option. In practice an individual (or household) purchases a bundle of servings (often 10 or more) in a single box. Instead of treating the problem as a discrete-continuous problem, we treat the package size as an additional characteristic, and use *servings* as the unit of observation. This is important because package sizes decline substantially over the course of the sample. Prior literature often aggregates different package sizes to form a composite *Honey Nut Cheerios* product.

We allow consumers to have heterogeneous preferences (random coefficients) over a number product characteristics. We allow for (potentially) correlated random coefficients on a constant (which captures variation in the taste for the outside option), on price p_{jt} and package size as well as for the principal components of the nutritional information $x_j = f(\tilde{x}_j)$. Furthermore, we allow the distribution of these random coefficients to depend on observable demographics of consumers which frequent each DMA-chain: *household income* and *presence of children* which we code as a dummy.

We consider the following specification for the linear component of utility:

$$\delta_{jct} = d_{jc} - \overline{\alpha} p_{jct} + \beta x_{jt} + \eta_{c,t} + \Delta\xi_{jt} \quad (10)$$

[14]The most common choice is something like the population of the three digit zipcode taken from the most recent Census or American Community Survey Data and fixed for a store over time, or aggregated over stores for the entire DMA. We avoid this definition for a few reasons: the first is that geographic market definitions can be problematic if shopping districts contain many retailers but few residents, the second is that geographic definitions often don't account for competing stores within the same area not observed in the dataset, and the third is that a temporally fixed marketsize implies that the outside good share may fluctuate substantially across time.

[15]We might worry that milk is a complement in the production function for bowls of cereal. Even if the production function were Leontief, it would still provide usable variation in the overall size of the market.

In our most flexible specification, we allow for $d_{j,c}$ chain × product specific intercepts which capture persistent preferences for products across chains as well as $\eta_{c,t}$ time or chain-time specific fixed effects, which capture variation in the taste for the outside good at the store-week level. In more restrictive specifications we consider d_j product fixed-effects which do not vary by chain or brand (*Honey Nut Cheerios* rather than brand-size *14 oz. Honey Nut Cheerios* specific fixed effects. Depending on the nature of the product fixed effects, the non-time varying component of x_{jt} may be subsumed into the fixed effect.

We estimate the parameters of the demand model following the approach of Berry et al. (1995). This means we need instruments which shift supply but not demand. The obvious choice for such instruments are exogenous cost shifters. One possibility (which we use) is to measure commodity prices of the main ingredient over time. This has advantages and disadvantages. The advantage is that commodity prices of corn are different from commodity prices of wheat and we can compare retail prices of corn-based cereals to those of wheat-based cereals. The disadvantage is that this instrument provides no geographic variation to explain prices in different stores for the same product at the same time. We plot those instruments in Figure 15.

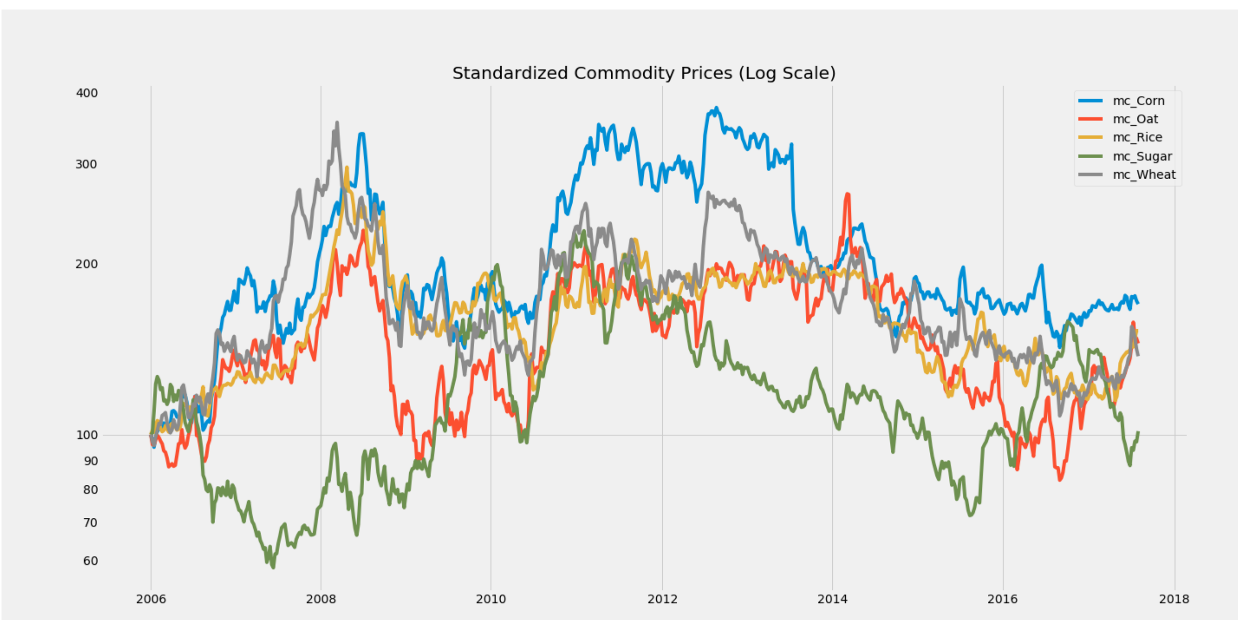


Figure 15: Input Prices

A second type of instrument are markup shifting instruments, which are often referred to as "BLP instruments". These instruments are based on the characteristics of other brands offered in the same market. As competing products become more similar to product j then we expect the markups of product j to fall. These instruments have the advantage that they vary both in the cross section and over time (as the set of products varies across stores and within a store over time). The main complaint about "BLP instruments" is that they can be weak (see Armstrong (2013)). Because the identification condition is a nonlinear conditional moment restriction $E[\Delta\xi_{jt}|X_t] = 0$ any (potentially nonlinear) function of X_t is a valid instrument $E[\Delta\xi_{jt} \cdot f(X_t)] = 0$. Ghandi and Houde (2016) propose "differentiation instruments" which amounts to counting up neighboring products in characteristic space. We construct differentiation instruments

A third type of instrument that is often employed are the so-called "Hausman instruments" or prices of the same product in other geographic markets. That is we could use the price of *Honey Nut Cheerios 14 oz.* in Boston as an instrument for the price of *Honey Nut Cheerios 14 oz.* in Chicago. More concretely, the purpose of Hasuman instruments is to measure cost shocks in the production process by using prices in several geographic markets at the same time. The idea being that if prices rise in many markets this could be due to a cost shock. The problem is that prices may rise in many markets for other reasons, including increased overlap in financial ownership of the manufacturers. If we believe that common ownership has an effect on prices, then the prices in other markets should no longer be valid instruments, however, under the null hypothesis of no common ownership effects these instruments might still be exogenous.

8.2 Supply

Our goal is to determine whether or not our data reflect a particular model of conduct, e.g. κ as opposed to some alternative, in particular κ^{MAX} or κ^{COL}. Fortunately, from Section 3, we have that κ is directly recoverable from data, and Section 6 develops clear implication for prices.

For notational purposes, let $O(\kappa)$ transform the firm-level matrix of cooperation into a product-level matrix of cooperation, where the i-jth entry corresponds to the f-gth entry in κ, with f the owner of product i and g the owner of product j. Recall that given estimates of demand $\mathbf{q}(\mathbf{p_t})$, its derivatives $\Omega(\mathbf{p_t})$ with elements $\Omega_{jk}(\mathbf{p_t}) = \frac{\partial q_j}{\partial p_k}$, and an ownership/conduct matrix $O(\kappa)$, we can write the $J \times J$ system of static-Betrand-Nash FOC's:

$$\begin{aligned} (\mathbf{p} - \mathbf{mc}) &= (O(\kappa) \odot \Omega(\mathbf{p}))^{-1} q(\mathbf{p}) \\ \rightarrow \mathbf{mc} &= \mathbf{p} - \underbrace{(O(\kappa) \odot \Omega(\mathbf{p}))^{-1} q(\mathbf{p})}_{\equiv MU(\kappa,\theta,\mathbf{p})} \end{aligned}$$

This lets us write:

$$mc_{jt} = p_{jt} - MU_{jt}(\kappa, \theta, \mathbf{p}) \tag{11}$$

From here, the literature gives us a few different approaches. The first is to specify a functional form for mc_{jt}, for example (Berry et al., 1995) employ $f(\cdot) = \log(\cdot)$:

$$f(mc_{jt}) = g(x_{jt}, z_{jt}) + \omega_{jt} \tag{12}$$

where x_{jt} are product characteristics (as in demand) and z_{jt} are excluded (from demand) cost-shifting instruments (such as the commodity prices of corn, rice, wheat, etc.).[16] If we plug (12) into (11), we obtain:

$$\omega_{jt} = f(p_{jt} - MU_{jt}(\kappa, \theta, \mathbf{p})) - g(x_{jt}, z_{jt}) \tag{13}$$

We can construct moment conditions of the form $E[\omega_{jt}|x_{jt}, z_{jt}, \eta_{jt}] = 0$ where ω_{jt} represents the unobserved component of marginal cost, x_{jt} are the usual (exogenous) product characteristics, and η_{jt} are excluded *demand shocks* which shift demand but do not effect marginal costs. Unlike exogenous cost shifters z_{jt} which are generally hard to come by, demand shifters η_{jt} are more readily available (for example: observable changes in store traffic, possibly temporary feature or displays in store- so long as the timing is determined by the retailer, etc.).

The first approach to testing for common ownership would include estimating demand and supply *simultaneously* by stacking the two sets of moment conditions $E[\Delta\xi_{jt}|x_{jt}, z_{jt}] = 0$ and $E[\omega_{jt}|x_{jt}, z_{jt}, \eta_{jt}] = 0$ and estimating via GMM while imposing a particular choice of $O(\kappa)$. The original (Berry et al., 1995) paper is an example of this approach where $O(\kappa)$ was given by the multiproduct oligopoly ownership matrix. Bresnahan (1987) estimates both supply and demand simultaneously and then tests non-nested hypotheses for the conduct which best matches the data.

[16]We do not include the BLP instruments here because they shift $MU_{jt}(\kappa, \theta, \mathbf{p})$

In a similar vein, Villas-Boas (2007) conducts a "horse-race" among various different choices for $O(\kappa)$ including double marginalization, vertical integration, etc. and essentially selects the model which best fits the data. Villas-Boas (2007) conducts a two step approach where $\hat{mc}_{jt}$ is recovered via (11) under an assumption of conduct $O(\kappa)$ and then in a second stage a regression of the form (12) is run.

All of the approaches assume that we have estimated a correctly specified model of demand. Furthermore, both of the aforementioned approaches also assume that we have a correctly specified model for marginal costs, and use the model of marginal costs to identify which set of markups is consistent with the data. Implicitly we require not only marginal cost shifters, but rather the true functional form for marginal costs in order to discriminate among alternatives for $O(\kappa)$.

We take a different approach in the spirit of Berry and Haile (2014) which relies on *exclusion restrictions* for (13). Under the null that $MU_{jt}(\kappa, \theta, \mathbf{p})$ is correctly specified, it should be the case η_{jt} can affect the (residual) demand for j and the price p_{jt} but only through the MU_{jt} term. However, it should not affect the residual component of marginal cost $mc_{jt} \equiv p_{jt} - MU_{jt}(\kappa, \theta, \mathbf{p})$. This suggests a different kind of test statistic similar to the test of over-identifying restrictions. We can extend (12) so that:[17]

$$f(mc_{jt}(\kappa)) = g(x_{jt}, z_{jt}) + \gamma \cdot \eta_{jt} + \omega_{jt}$$

Under the null that $MU_{jt}(\kappa, \theta, \mathbf{p})$ is correctly specified it should be that $\gamma = 0$. Under an alternative (and incorrect) hypothesis regarding conduct $O(\kappa)$ we could have that $\gamma \neq 0$. Under the null, we could construct:

$$\omega_{jt}(\kappa) \equiv f(mc_{jt}(\kappa)) - g(x_{jt}, z_{jt})$$

From here we can directly impose: $E[\omega_{jt}|x_{jt}, z_{jt}, \eta_{jt}] = 0$ and estimate the parameters of $g(\cdot)$ via IV-GMM or GEL.[18] A test statistic of this form has the advantage that it merely relies on an exclusion restriction and does not require that we stipulate the correct functional form for $f(mc_{jt}) = g(x_{jt}, \gamma)$ just the weaker condition that $E[\omega_{jt}\eta_{jt}|(x_{jt}, z_{jt})] = 0$. The downside is that we may still have difficulty distinguishing among two conduct models which yield similar markups.

[17] Here $g(\cdot)$ stands in for some semi-parametric function of (x_{jt}, z_{jt}) such as a sieve or series estimator.
[18] We provide more detail in Appendix A.4.

The other challenge for our approach is that it relies on over-identifying restrictions. We require variables η_{jt} which shift the (residual) demand for j without affecting marginal costs mc_{jt}. A potential source of exclusion restrictions are cost shifters for competing products. If we have the correct form for $MU_{jt}(\kappa, \theta, \mathbf{p})$ then cost shocks to competing goods should not affect $c_{jt} \equiv p_{jt} - MU_{jt}(\kappa, \theta, \mathbf{p})$. (A cost shock to k affects p_{kt} which in turn might affect demand for j and MU_{jt} but should not affect mc_{jt} directly.)[19] In our case this means we can use the price of corn as an exclusion restriction for the price of *Rice Krispies* and the price of rice as an exclusion restriction for the price of *Corn Flakes*. If we had additional information on MC_{jt} (for example margins from accounting data)[20] we could impose that as an additional restriction.

An additional set of exclusion restrictions arises when there are observable changes in κ, for example, from a financial event such as the Blackrock acquisition of BGI. this should produce large observable changes in $MU_{jt}(\kappa, \theta, \mathbf{p})$. We expect these restrictions to be quite powerful in discriminating among different models of conduct. Under the true model, κ should be uncorrelated with marginal costs, but under an incorrectly specified model it may.

We might also expect that absent any changes in x_{jt} or z_{jt} between period t and t' marginal costs before and after the event should be similar. That is, the wrong model will perceive changes in κ as large changes in ω_{jt} while the correct model should produce a smooth function of $mc_{jt}|(x_{jt}, z_{jt})$.

9 Pricing, Costs, and Implications of Conduct Models

The results we present in this section rely on a simplified version of our full demand model. To begin we consider a nested logit demand with all products entering the same nest and only the outside option in the other nest. As in Berry (1994), this yields the following estimating equation:

$$\ln s_{jmt} - \ln s_{0mt} = \gamma_{jm} + \beta x_{jmt} - \alpha p_{jmt} + \sigma \ln s_{j|gmt} + \varepsilon_{jmt} \tag{14}$$

[19] An alternative might be temporary feature or displays if they are chosen by the retailer and retail prices are chosen by the manufacturer (or vice versa).

[20] an example of this approach (Crawford et al., 2018).

There are two endogenous variables to instrument for. The first is the price p_{jmt} and the second is the $\ln s_{j|gmt}$ (within-group share). One challenge to any potential instrumental variables strategy is that if we include too many fixed effects we may not have sufficient residual variation in the instrument. For example, if we allow γ_{jm} $Prod \times Chain$ FE then the residual variation is within a chain-product combination across time. A common instrument for σ is the (log) number of products available in that chain-week. We report the estimates of the first stage in Table 11. We report the t-statistics in brackets. We find that the number of products is a strong instrument for σ even after including $Prod \times Chain$ FE. Though not reported, we find that neither the number of products, nor the "BLP instruments" are strong instruments for price (α) once we include the $Prod \times Chain$ FE even though they are quite strong without these FE.[21] The problem is that the residual variation in price at the product-chain level comes largely though the timing of sales for which we do not have good predictors. The R^2 of the first-stage regressions of price on just the fixed effects is large $R^2 = 0.89$. The estimates we present include instruments for σ but do not instrument for α as weak instruments are often worse than no instruments at all [CITE: Stock and ?].

Table 11: First Stage for σ

log(# prods)	-0.645	-0.941	-1.016	-0.784	-0.873
(for σ)	[-309.58]	[-254.36]	[-259.44]	[-217.82]	[-225.71]
Prod FE	3,238	3,238	3,238	n/a	n/a
Chain FE	no	36	36	n/a	n/a
Week FE	no	no	467	no	467
Prod X Chain	no	no	no	30,171	30,171
R^2 σ	0.570	0.598	0.604	0.706	0.712
R^2 p_{jt}	0.837	0.855	0.857	0.891	0.892

We estimate the nested logit model from (14) and report the results in Table 12. Rather than reporting standard errors in parenthesis we report t-statistics. Because all parameters are scaled by $(1-\sigma)$, for convenience we report the adjusted price sensitivity $\alpha/(1-\sigma)$.

[21] Even the quasi-optimal Ghandi and Houde (2016) instruments yield an F-stat < 2 for price in the first stage.

Table 12: Nested Logit Estimates

	(1)	(2)	(3)	(4)	(5)
α	-12.52	-3.690	-3.373	-8.735	-6.645
	(-39.05)	(-13.82)	(-5.91)	(-20.82)	(-8.56)
σ	0.284	0.686	0.705	0.581	0.553
	(27.91)	(117.26)	(98.47)	(60.61)	(43.22)
unemp $\times$ branded	0.0155	0.0212	-0.00677	0.0256	-0.00504
	(16.20)	(51.97)	(-12.59)	(41.22)	(-6.49)
$\alpha/(1-\sigma)$	-17.48	-11.76	-11.43	-20.84	-14.87
	(-26.15)	(-16.14)	(-6.50)	(-26.18)	(-8.79)
Product FE	Yes	Yes	Yes	Yes	Yes
Chain FE	No	Yes	Yes	No	No
Chain $\times$ Product FE	No	No	No	Yes	Yes
Week FE	No	No	Yes	No	Yes
N	997329	997329	997329	995421	995421
Adj R^2	0.6610	0.9388	0.9503	0.9042	0.9237

t statistics in parentheses.

α, σ instrumented with number of products, and own ingredient commodity prices.

Consider the multi-product Bertrand FOCs:

$$\begin{aligned}\arg\max_{p\in\mathcal{J}_f} \pi_f(\mathbf{p}) &= \sum_{j\in\mathcal{J}_f}(p_j - c_j)\cdot q_j(\mathbf{p}) + \kappa_{fg}\sum_{k\in\mathcal{J}_g}(p_k - c_k)\cdot q_k(\mathbf{p}) \\ 0 &= q_j(\mathbf{p}) + \sum_{k\in\mathcal{J}_f}(p_k - c_k)\frac{\partial q_k}{\partial p_j}(\mathbf{p}) + \kappa_{fg}\sum_{l\in\mathcal{J}_g}(p_l - c_l)\frac{\partial q_l}{\partial p_j}(\mathbf{p})\end{aligned}$$

It is helpful to define the matrix $\Omega_{(j,k)}(\mathbf{p}) = -\frac{\partial q_j}{\partial p_k}(\mathbf{p})$:

$$O(\kappa)_{(j,k)} = \left\{\begin{array}{ll} 1 & \text{for } j\in\mathcal{J}_f \\ \kappa_{fg} & \text{for } j\in\mathcal{J}_f, k\in\mathcal{J}_g \\ 0 & \text{o.w} \end{array}\right\}$$

We can re-write the FOC in matrix form:

$$q(\mathbf{p}) = (O(\kappa)\odot\Omega(\mathbf{p}))\cdot(\mathbf{p}-\mathbf{mc})$$

Recovering PCM is easy:

$$\begin{aligned}(\mathbf{p}-\mathbf{mc}) &= (O(\kappa)\odot\Omega(\mathbf{p}))^{-1}q(\mathbf{p}) \\ \rightarrow \mathbf{mc} &= \mathbf{p} + (O(\kappa)\odot\Omega(\mathbf{p}))^{-1}q(\mathbf{p})\end{aligned}$$

We can recover marginal costs under the different assumptions regarding κ_{fg}. We can assume: (1) no common ownership $\kappa_{fg} = 0$; (2) that κ_{fg} follows the O'Brien and Salop (2000) weights; (3) or the Crawford et al. (2018) weights. We can plot implied marginal costs in Figures 16-17. Accounting estimates imply that we expect markups in the 40-60% range which are most consistent with the $\kappa = 0$ estimates as prices are around 23-25 cents per serving for most products (except for Quaker Oats' less expensive puffed products with an average price of 19 cents). If firms were incorporating κ when setting prices then implied marginal costs would be much too low, particularly for the O'Brien and Salop (2000) weights.

To explore the mechanism behind the common ownership κ's affect on prices consider Table 13. In this table, we see that absent any common ownership effect $\kappa = 0$, that Kellogg's and General Mills recover around 25% of sales when they increase prices of an average product while Quaker Oats recovers around 9% and Post around 12%, with around 20% of customers choosing the no-purchase option. If the four major firms fully colluded, or were allowed to

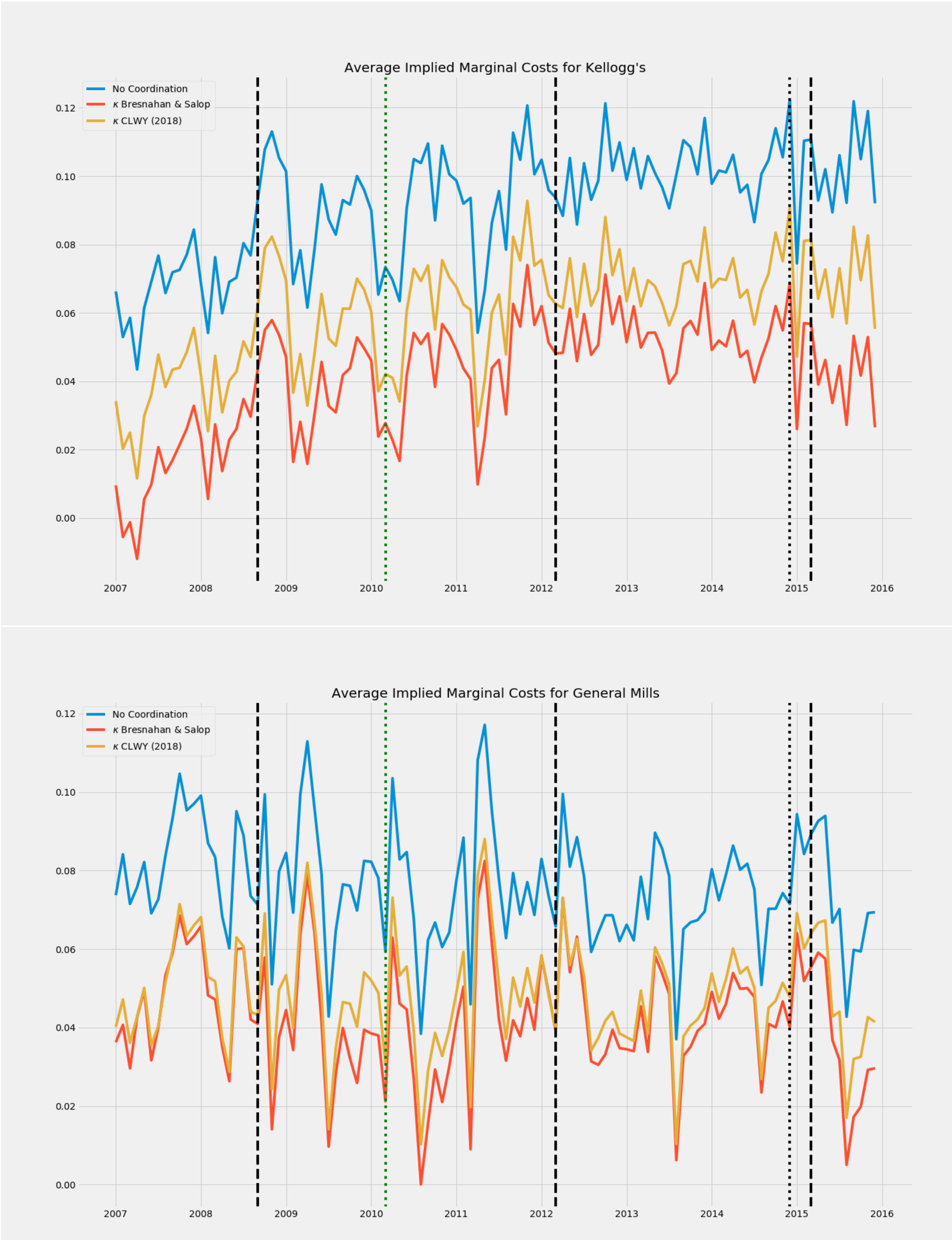


Figure 16: Implied Marginal Cost

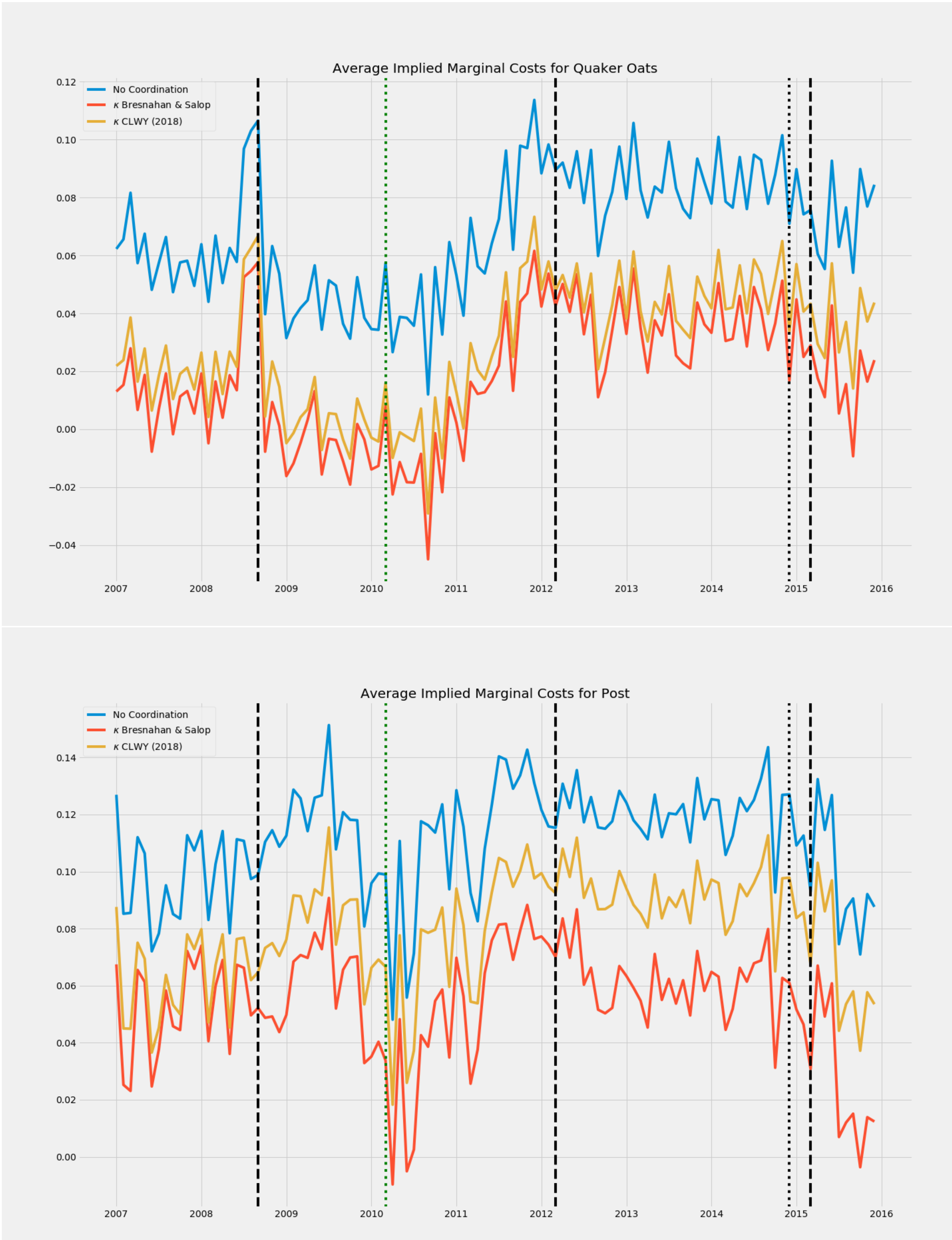


Figure 17: Implied Marginal Cost

	$\kappa = 0$	κ_{OBS}	κ_{CLWY}	$\kappa = 1$	Outside Good
K	24.79	47.96	38.23	65.71	19.57
GIS	25.83	39.23	36.55	65.67	19.66
PEP	9.18	32.57	26.79	64.85	19.59
POST	11.73	43.44	26.67	64.77	19.43
Market	18.57	35.68	29.36	54.76	19.53

Table 13: Effective Diversion Ratios Under Common Ownership

	$\kappa = 0$	κ_{OBS}	κ_{CLWY}	Actual Price
K	9.98	4.08	6.76	25.72
GIS	7.19	3.55	4.59	23.21
PEP	6.89	1.63	3.35	19.91
POST	9.57	2.64	6.52	22.79
Market	7.79	3.47	5.26	22.55

Table 14: Implied Marginal Costs (Cents per serving)

merge $\kappa = 1$ the fraction of diverted sales recaptured would increase to around 65%. Under common ownership we see a larger effective fraction of diverted sales recaptured (from an overall average of around 19% to an average of 30-35%), which would represent a substantial softening of competition.

We can see that the common ownership κ weights imply substantially lower marginal costs Table 14 and larger markups Table 15. Particularly under the O'Brien and Salop weights we see markups that are 84% or larger, and 72% or larger under the Crawford, et. al weights.

	$\kappa = 0$	κ_{OBS}	κ_{CLWY}	Actual Price
K	61.19	84.12	73.72	25.72
GIS	69.01	84.70	80.21	23.21
PEP	65.40	91.80	83.15	19.91
POST	58.03	88.40	71.39	22.79
Market	65.44	84.62	76.68	22.55

Table 15: Implied Markups (Percent)

	GIS-K	GIS-PEP	GIS-POST	K-PEP	K-POST	PEP-POST	$\kappa = 1$	κ_{OBS}	κ_{CLWY}	p-obs
K	16.69	0.37	0.64	0.45	0.79	0.13	41.70	19.49	10.95	25.72
GIS	2.44	0.58	1.04	0.41	0.71	0.16	45.58	14.85	10.31	23.21
PEP	0.55	18.73	0.25	17.26	0.22	0.24	63.07	24.31	16.38	19.91
POST	0.68	0.17	17.18	0.16	15.92	3.67	54.63	27.20	12.19	22.79
Market	6.95	1.73	2.94	1.57	2.68	0.57	43.20	17.40	10.25	22.55

Table 16: Counterfactual Prices and Merger Comparisons (Percentage Price Change)

As a final exercise we compare the effects of common ownership to all possible pairwise mergers between the four largest firms. Here we find that even a pairwise merger between two 30% marketshare firms (Kellogg's and General Mills) would only increase the overall cereal price index by 7% while moving from a world with no common owner incentives to one with common owner incentives as described by the OBS weights would increase prices by 17.4% and 10% under the CLWY weights. The effects are quite substantial though smaller than full collusion (a 43% price increase).

10 Conclusion

We have showed that common ownership has increased over time in the US economy, and that this is associated with an increase in implied profit weights that firms should place on other firms across the economy, but also within industries. While a "reduced form" approach to analyzing this problem seems superficially attractive, it is flawed. Instead, a structural approach to the problem looking at specific industries is the correct approach. Examining the ready-to-eat cereal market, we find that classic competition is most consistent with demand and supply data, but that the implied effects of common ownership are large relative to mergers that would be clearly blocked out of antitrust concerns.

References

Anton, M., Ederer, F., Gine, M., and Schmalz, M. C. (2016). Common ownership, competition, and top management incentives. Ross School of Business Paper No. 1328.

Armstrong, T. (2013). Large market asymptotics for differentiated product demand estimators with economic models of supply. Working Paper.

Azar, J., Raina, S., and Schmalz, M. C. (2016). Ultimate ownership and bank competition. Working Paper.

Azar, J., Schmalz, M. C., and Tecu, I. (2017). Anti-competitive effects of common ownership. Ross School of Business Paper No. 1235.

Backus, M., Conlon, C., and Sinkinson, M. (2018). The common ownership hypothesis: Theory and evidence. Brookings Working Paper.

Ben-David, I., Franzoni, F., Moussawi, R., and Sedunov, J. (2018). The granular nature of large institutional investors. Swiss Finance Institute Research Paper Series.

Berle, A. and Means, G. (1932). *The Modern Corporation and Private Property*. MacMillan.

Berry, S., Levinsohn, J., and Pakes, A. (1995). Automobile prices in market equilibrium. 63(4):841–890.

Berry, S. T. (1994). Estimating discrete-choice models of product differentiation. *The RAND Journal of Economics*, 25(2):242–262.

Berry, S. T. and Haile, P. A. (2014). Identification in differentiated products markets using market level data. *Econometrica*, 82(5):1749–1797.

Bresnahan, T. F. (1982). The oligopoly solution is identified. *Economics Letters*, 10(1-2):87–92.

Bresnahan, T. F. (1987). Competition and collusion in the american automobile industry: The 1955 price war. *The Journal of Industrial Economics*, 35(4):457–482.

Bresnahan, T. F. and Salop, S. C. (1986). Quantifying the competitive effects of production joint ventures. *International Journal of Industrial Organizatoin*, 4:155–175.

Crawford, G. S., Lee, R. S., Whinston, M. D., and Yurukoglu, A. (2018). The welfare effects of vertical integration in multichannel television markets. *Econometrica*.

DellaVigna, S. and Gentzkow, M. (2017). Uniform pricing in us retail chains. Technical report, NBER 23996.

Demsetz, H. (1973). Industry structure, market rivalry, and public policy. *Journal of Law and Economics*, 16(1):1–9.

Elhauge, E. (2016). Horizontal shareholding. *Harvard Law Review*, 109:1267–1317.

Friedman, M. (1953). *Essays in Positive Economics*, chapter The Methodology of Positive Economics. University of Chicago Press. (unfinished citation).

Ghandi, A. and Houde, J.-F. (2016). Measuring substitution patterns in differentiated products industries. Working Paper.

Gramlich, J. and Grundl, S. (2017). Estimating the competitive effects of common ownership. FEDS Working Paper 2017-029.

Hong, H., Preston, B., and Shum, M. (2003). Generalized empirical likelihood based model selection criteria for moment condition models. *Econometric Theory*, 19(06):923–943.

Johnson, S., Porta, R. L., Lopez-De-Silanes, F., and Schleifer, A. (2000). Tunneling. *American Economic Review: Papers and Proceedings*, 90(2):22–27.

Kennedy, P., O'Brien, D. P., Song, M., and Waehrer, K. (2017). The competitive effects of common ownership: Economic foundations and empirical evidence. Working Paper.

López, A. and Vives, X. (2018). Overlapping ownership, r&d spillovers, and antitrust policy. Working Paper.

Miller, N. and Weinberg, M. (2017). Understanding the price effects of the millercoors joint venture. *Econometrica*, 85(6):1763–1791.

Nevo, A. (2000). Mergers with differentiated products: The case of the ready-to-eat cereal industry. *RAND Journal of Economics*, 31(3):395–421.

Nevo, A. (2001). Measuring market power in the ready-to-eat cereal industry. *Econometrica*, 69(2):307–342.

O'Brien, D. and Salop, S. (2000). Competitive effects of partial ownership: Financial interest and corporate control. *Antitrust Law Journal*, 67(3):559–614.

O'Brien, D. P. (2017). Price-concentration analysis: Ending the myth, and moving forward. Working Paper.

Porta, R. L., Lopez-De-Silanes, F., and Schleifer, A. (1999). Corporate ownership around the world. *The Journal of Finance*, 54(2):471–517.

Posner, E. A., Scott-Morton, F. M., and Weyl, E. G. (2017). A proposal to limit the anti-competitive power of institutional investors. *Antitrust Law Journal*, 81(3).

Rotemberg, J. (1984). Financial transactions costs and industrial performance. Working Paper.

Villas-Boas, S. B. (2007). Vertical relationships between manufacturers and retailers: Inference with limited data. *The Review of Economic Studies*, 74(2):625–652.

Appendices

A Product Market Data Appendix

A.1 Product Characteristics

Because we are not interested in the relationship between demand and the 19 nutritional characteristics per se, and because the characteristics are often correlated with one another. For example total calories are a linear function of fat, carbohydrates, and protein; or sugar and dietary fiber are a source of carbohydrates.We report average nutritional information on a per-serving basis in Table A-1 We use a principal components approach to reduce the dimension of the product space and provide a lower dimensional basis of *orthogonal* factors. We describe our procedure in detail below.

Consider a $(J \times K)$ matrix X with rows denoting products and columns denoting nutritional characteristics. Principal components constructs an eigen-decomposition of the covariance matrix $(X'X) = W\Lambda W'$ where W is the matrix of eigenvectors and Λ is a diagonal matrix of eigenvalues. We can approximate the covariance matrix by considering the largest $k < K$ eigenvalues and their corresponding eigenvectors and construct a basis in dimension k rather than K. This yields the approximation $(X'X) \approx W\Lambda_k W'$ where Λ_k is the same diagonal matrix as Λ but with zeros for all eigenvalues except the k largest. This lets us write $F = XW_k'$ where W_k is the $J \times k$ matrix of eigenvectors which correspond to the k largest eigenvalues in Λ. F is an orthonormal approximation in dimension k to the basis spanned by the covariance matrix $(X'X)$ in dimension K.

This has some advantages: for one we have reduced the number of dimensions of nutritional information that we need to keep track of. This is especially helpful when we want consumers to have heterogeneous preferences (random coefficients) across these characteristics. The second advantage is that the columns of F are now orthogonal. In general this makes estimating parameters much easier because we no longer have to worry about columns of X being highly correlated with one another (such as sugar and carbohydrate content). The disadvantage is that the transformed components of F can be difficult to interpret.

	mean	standard deviation
calories	133.636	43.111
calories_from_fat	12.785	11.659
vitamin_a_dv	7.715	6.637
vitamin_c_dv	9.910	15.105
iron_dv	30.718	24.454
calcium_dv	4.844	10.456
serving_weight (g)	36.111	11.604
As a Percentage of Weight		
cholesterol	0.080	0.515
dietary_fiber	0.075	0.053
monounsaturated_fat	0.007	0.015
polyunsaturated_fat	0.006	0.012
protein	0.077	0.042
saturated_fat	0.006	0.015
sodium $\times 1000$	4.403	2.186
sugars	0.264	0.125
total_carbohydrate	0.811	0.076
total_fat	0.041	0.035
other_carbs	0.472	0.114

Table A-1: Summary Statistics of Product Characteristics (3478 Products)
% Coverage reports share of Nielsen's calculated All Commodity Volume (ACV) by DMA-Channel.

When we apply this to our dataset, instead of working with the covariance matrix $(X'X)$ we work with the weighted covariance matrix $(X'\Omega X)$ where Ω is an $J \times J$ diagonal matrix of weights where the weights correspond to the brand-level marketshares for each row of X. This prevents less popular brands with strange product characteristics from becoming outliers and distorting our transformed product space. The first five components comprise around 75% of the overall variation of the 19 nutritional characteristics.

[UPDATE]We display the first two components of our transformed product space in Figure A-1. Those brand with values of component one tend to be more dense cereals (higher weight, smaller serving sizes, higher protein and fat content, often granola or museli based cereals) whereas those with lower values of component one tend to be less dense cereals ("puffs", "crunch") and those in the middle are things like "bran flakes" or "O's". The second component is a bit harder to interpret, but products with low values of component 2 tend to be candy-like children's cereals (Reese's Puffs, Herhsey's Cookies & Cream, Frosted Toast Crunch, etc.). The third component (not displayed) appears to mostly be picking up vitamin content with the all of the low values corresponding to General Mills *Total* brands which are essentially a multivitamin in cereal form.

A.2 Marketsize

Recall that our market is defined as a DMA-chain-week. In other words we combine weekly sales for stores within the same chain, so long as those chains belong to the same Nielsen DMA. Our market definition is anyone who walks into one of that chains stores that week. We also know from the Nielsen Panelist data that RTE cereal is purchased by consumers on 27.6% shopping trips. Since we don't observe the number of consumers to walk into a store in a particular week, we try to measure it using two other commonly purchased categories: milk and eggs. We run the following fixed effects regression:

$$\log q_{c,t}^{RTEServings} = \beta_1 \log q_{c,t}^{milk} + \beta_2 \log q_{c,t}^{eggs} + \gamma_c + \varepsilon_{ct}$$

We predict the weekly number of RTE Cereal servings purchased as a function of milk volume (in gallons) and egg purchases (in dozens) along with fixed effects at the DMA-Chain level. We don't include any time varying fixed effects because those might be correlated with the overall price level for RTE cereal. Instead we rely on the milk and eggs purchases to capture the time varying traffic of individual chains. We then define our marketsize as, where λ is a

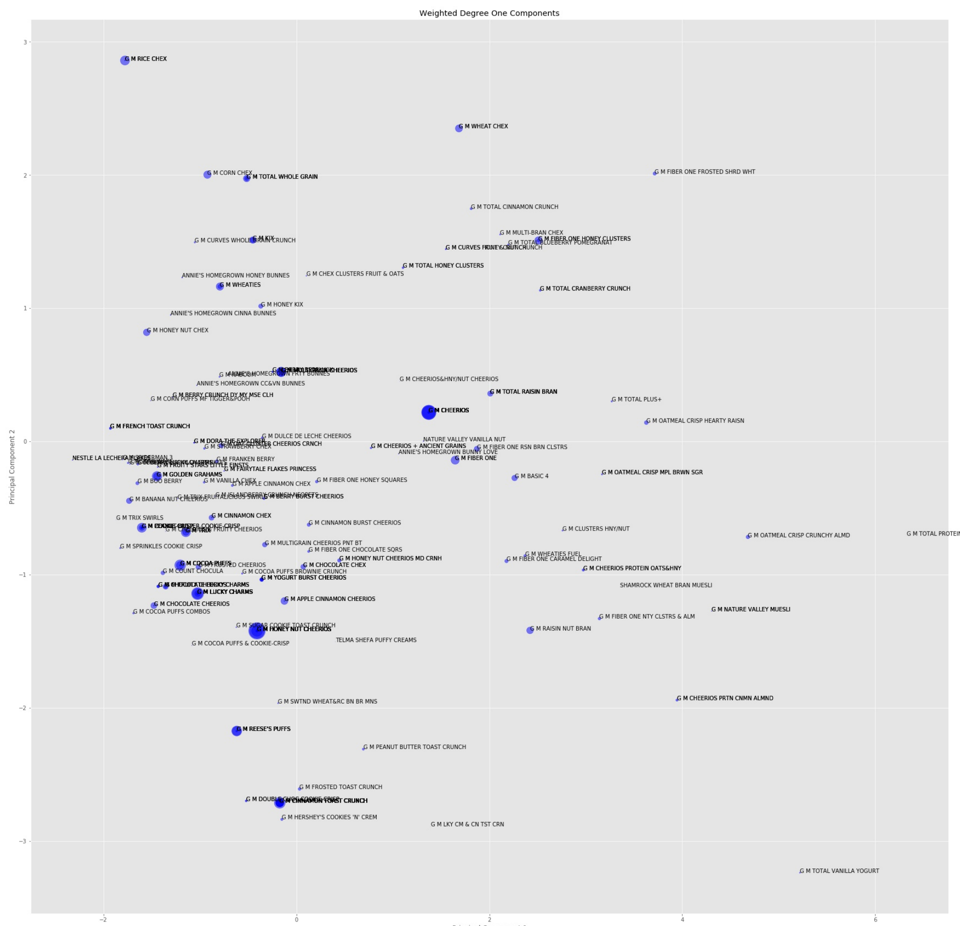


Figure A-1: Principal Components of Nutritional Data: Top 250 Brands
Dot size indicates brand-level marketshares

constant chosen to match the (average) outside good share of $s_0 = 1 - 0.276 = 0.723$:

$$\hat{M}_{c,t} = \lambda \cdot \left(\hat{\beta}_1 \log q_{c,t}^{milk} + \hat{\beta}_2 \log q_{c,t}^{eggs} + \hat{\gamma}_c\right)$$

We report the results of those regressions in Table A-2. We find that on its own 1% increase in milk sales increases the sales of RTE Cereal (measured in servings) by 0.88% while a 1% increase in egg purchases increases cereal purchases by 0.81%. While not displayed in the table, chain fixed effects alone explain around 94% of the variation in the cereal sales, while the addition of milk and egg sales increase the R^2 to almost 99% and the within $R^2 = 0.7742$.

VARIABLES	(1) Log Servings	(2) Log Servings	(3) Log Servings
milk_vol	0.886***		0.535***
	(0.0666)		(0.0572)
egg_vol		0.809***	0.407***
		(0.0587)	(0.0648)
Observations	21,836	21,833	21,833
R-squared	0.983	0.982	0.987
Chain+Week FE	YES	YES	YES

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table A-2: Distribution of Inside Good Shares by DMA-Chain-Week

In this process we find a very small number of outliers and drop 249 out of 21,587 observations (1.1% of the sample by observations, 0.8% by revenue) by trimming the largest and smallest values by the outside good share. We report the resulting distribution of the inside good share $1 - s_{0ct} = \frac{\sum_j s_{jct}}{M_{ct}}$ in Figure A-2, which we have standardized around the 27.6% purchase probability.

This is a bit different from the usual approach. Most papers, including Nevo (2001), construct marketsize using population demographics for the surrounding area (the population of the DMA, the population of the zipcode, etc.). That approach generally does not capture the fact that some stores face more competition than other stores (ie: we often don't know if there is a major competitor nearby or not). One disadvantage of our approach is that it can become potentially contaminated if sales of milk or eggs respond to the price for cereal or if consumers go to the supermarket just to buy cereal. Even if cereal and milk are complements (or Leontief in breakfast production) as long as the price of cereal doesn't causally impact

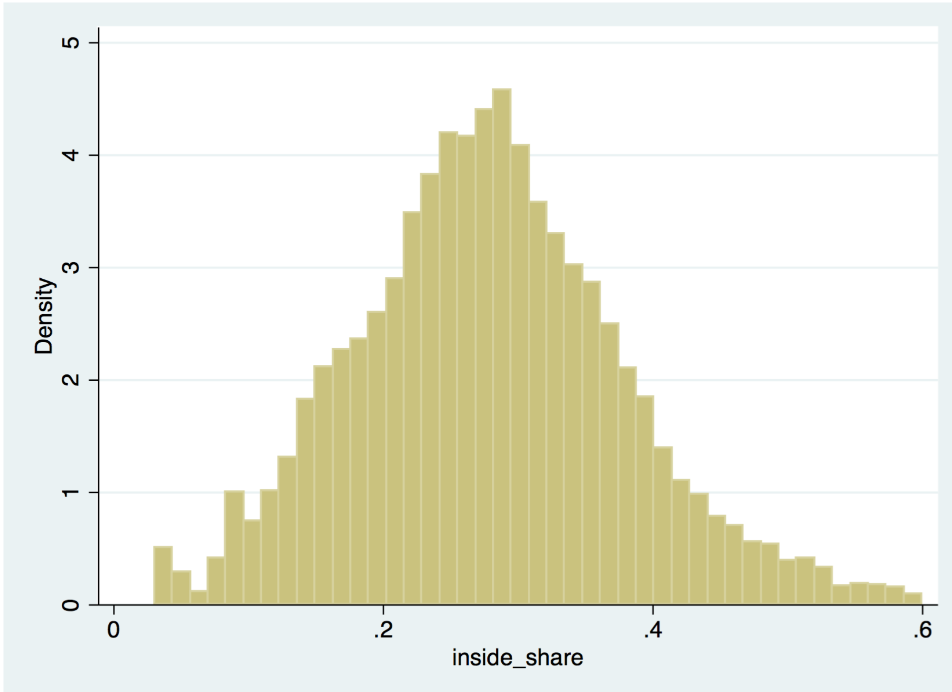


Figure A-2: Distribution of Inside Good Shares

the sales of milk and eggs we should be alright.

A.3 Chain Pricing

Following **??**, we find that for the most part prices do not vary across chains. We provide a plot in Figure A-3 in which each pixel represents a store-week price for a single UPC within a large chain. Each row tracks prices within a single store over time. We arrange the rows by their three-digit zipcode so that stores closer together in physical space are closer together within the matrix. Each column represents a weekly observation. A dark vertical bar represents a temporary sale. We can see visually from the heatmap plot that within the chain stores have very common prices and even time temporary sales simultaneously. There is some geographic dispersion. The colors in the top and bottom half of the graph are somewhat different from one another. This split corresponds to two separate DMAs (Denver and Phoenix) with prices being slightly higher in the Phoenix DMA than the Denver DMA. The white pixels are store-weeks where we are missing data.

Redacted

Figure A-3: Chain Pricing Heatmap

A.4 Testing

We follow the testing procedure in Hong et al. (2003). Given a set of moment restrictions $E[g(y_i, x_i, z_i, \theta)] = 0$. For example in linear IV $E[(y_i - \beta x_i)[x_i z_i]] = 0$ we can estimate a set of coefficients β which best satisfy the moment conditions. We consider the version of this problem based on the Generalized Empirical Likelihood (GEL) estimators. This estimator imposes that moment conditions hold exactly, but that the data are *resampled* in order to do so. In other words, each observation as assigned a weight π_i with the vector of weights living in the unit simples $\pi \in \Delta$. Estimation involves a search over both these weights π_i and the parameters governing the moment conditions θ.

$$\min_{\pi,\theta} \sum f(\pi_i) \quad \text{s.t.} \sum_i \pi_i \cdot g_i(y_i, x_i, z_i, \theta) = 0, \quad \pi \in \Delta.$$

There are different choices for $f(\pi_i)$ which correspond to different estimators, for example: Empirical Likelihood (EL): $f = -\log \pi_i$, Expontential Tilt (ET): $f = \pi \log \pi$, and Continuously Updated GMM (CUE): $f = \pi_i^2$. All of these estimators have the property that if the moment conditions are trivially satisfied then $\widehat{\pi}_i = \frac{1}{N}$ which we think about as the empirically observed weights. We can view each choice of f as a different way to penalize deviations from the $\frac{1}{N}$ weights.

The easiest estimator to work with is the EL estimator because it provides likelihood units.[22] We can consider the ELR test statistic which is just $ELR = (\sum_i \widehat{\pi}_i) - N \cdot \log N$. We can now use likelihood ratio style test statistics. This has the advantage that unlike say the pairwise Cox testing procedure used in Villas-Boas (2007) we can test and rank all models at once by their ELR statistics rather than using the optimal weighting matrix associated with model A to test model B and the optimal weighting matrix associated with model B to test model A (which could in theory lead to each model rejecting the other). We can also compare optimized CUE objective functions if we wanted something more in line with traditional GMM estimators but without the problem of multiple optimal weighting matrices.

[22]This is often called a non-parametric maximum likelihood estimator NPMLE. See (?).

B Ownership Data Appendix

Common ownership weights (κ's) are constructed from ownership data, which is obtained from data on investor holdings obtained from the Wharton Research Data Services (WRDS). The primary source is the s12 database of 13f SEC filings, which are required for institutional investment managers with at least $100 million dollars under management – this, for our purposes, is the universe of institutional investors.

One limitation of the 13f filings is that they double-count "short" positions. Both the borrower of the share and the lender count the it in their holdings for the purposes of the 13f, and therefore when the short interest is large, e.g. when the market anticipates trouble or potential bankruptcy, it may appear that institutional ownership of a firm exceeds 100%. This is measurement error attributable the double-counting of short interest.

To this end we merge in short interest tallies from Compustat, also available through WRDS. However, this data is at the firm-quarter level, not the investor-firm-quarter level, and therefore we cannot correct for measurement error in the common ownership terms. Instead, we use this data to correct total investor shares and retail shares by subtracting the short interest from the total investor share, which is the sum of holdings according to 13f filings.

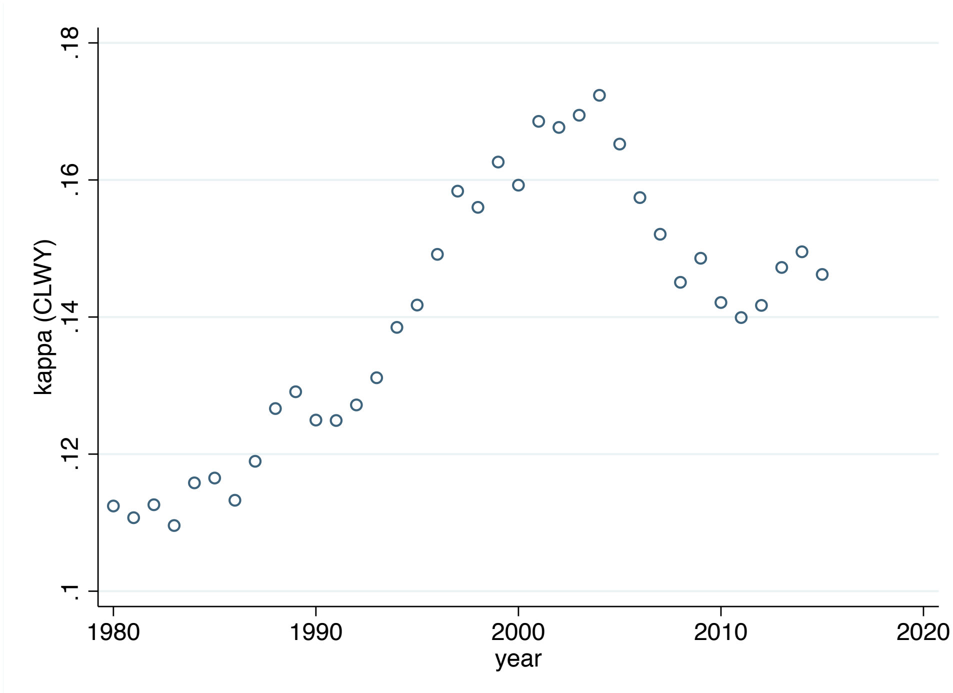
A.1 CLWY Weights

The alternative κ formulation proposed in CLWY for the vertical relations context can also be used to compute profit weights for the S&P 500. Doing so yields a qualitatively different result: the overall levels are much lower, and there is a reversal in trend towards the end of the sample.

A.2 S&P 500 Ownership Data

For the purposes of Section 4 we compute common ownership weights for all component firms in the S&P 500 from 1980 to 2016. These weights are computed from the same 13f filings as the cereal data. The universe of S&P 500 firms for each quarter is obtained from Capital IQ, also via WRDS, and so we are able to track entry and exit. Many firms appear

Figure A-4: Common Ownership over Time



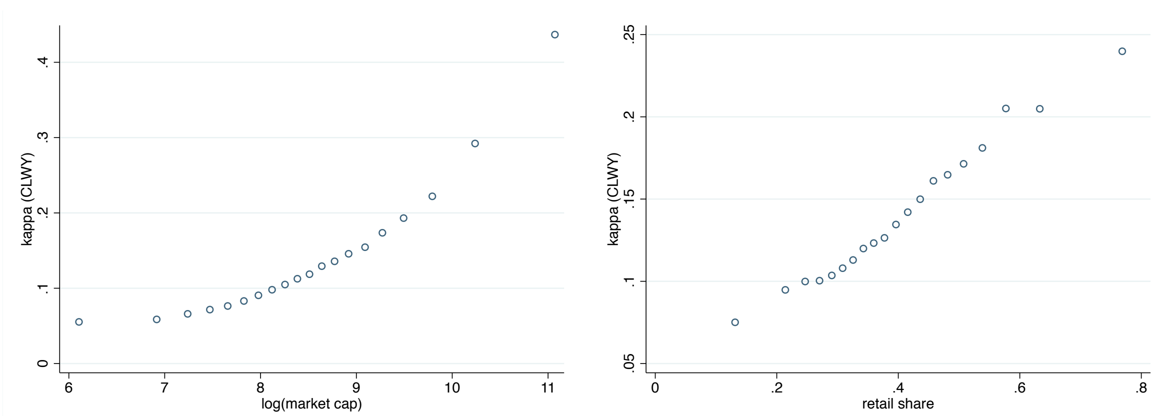


Figure A-5: Heterogeneity in Common Ownership, CLWY Weights

more than once in the index, especially airlines, which exit during periods of bankruptcy but then re-join. We exclude a small handful fo firms that are recorded as entering and exiting in the same period, as from manual inspection these appear to be database errors.

As discussed in the main text, we exclude firms that issue dual-class shares, because these explicitly divorce control rights and cash-flow rights in a way that render common ownership weights uninterpretable. Often they are used to consolidate control in the hands of founders

Table A-3: Correlations with κ (with segments)

	OS			CLWY		
	(1)	(2)	(3)	(4)	(5)	(6)
No. Segments	-0.0004	-0.0001	-0.0007	0.0025*	0.0023	0.0035*
	(0.0019)	(0.0019)	(0.0018)	(0.0012)	(0.0013)	(0.0013)
Year	0.0099*		0.0104*	-0.0023*		-0.0006
	(0.0004)		(0.0005)	(0.0003)		(0.0003)
Market Cap (in logs)	0.0924*	0.0929*	0.0837*	0.0658*	0.0664*	0.0515*
	(0.0036)	(0.0036)	(0.0041)	(0.0034)	(0.0034)	(0.0029)
Retail Share	0.6880*	0.7176*	0.6161*	0.2277*	0.2463*	0.1983*
	(0.0232)	(0.0240)	(0.0252)	(0.0159)	(0.0169)	(0.0140)
Mining, Quarrying, and Extraction	-0.0015	-0.0032		-0.0041	-0.0042	
	(0.0142)	(0.0142)		(0.0084)	(0.0085)	
Utilities	0.0125	0.0066		0.0316*	0.0285*	
	(0.0127)	(0.0126)		(0.0086)	(0.0086)	
Manufacturing	-0.0036	-0.0052		0.0061	0.0058	
	(0.0080)	(0.0080)		(0.0051)	(0.0052)	
Retail	-0.0129	-0.0149		-0.0030	-0.0040	
	(0.0129)	(0.0129)		(0.0080)	(0.0081)	
Information	0.0131	0.0100		0.0306*	0.0290*	
	(0.0145)	(0.0142)		(0.0143)	(0.0141)	
Finance and Insurance	-0.0294*	-0.0309*		-0.0213*	-0.0219*	
	(0.0132)	(0.0133)		(0.0071)	(0.0072)	
Other Sectors	0.0000	0.0000		0.0000	0.0000	
	(.)	(.)		(.)	(.)	
Year FE		✓			✓	
Firm FE			✓			✓
N	26638970	26638970	26638970	26638970	26638970	26638970

or management. To identify these stocks we used the Council of Institutional Investor's 2017 list of dual-class stocks[23]

[23] This is available at https://www.cii.org/files/3_17_17_List_of_DC_for_Website(1).pdf.

Table A-4: Correlations with κ (with diversification dummies)

	OS			CLWY		
	(1)	(2)	(3)	(4)	(5)	(6)
Diversified Firm (2+)	0.0027	0.0037	-0.0021	0.0062	0.0063	0.0066
	(0.0063)	(0.0063)	(0.0071)	(0.0042)	(0.0043)	(0.0047)
Highly Diversified Firm (5+)	0.0006	-0.0001	-0.0014	0.0053	0.0041	0.0071*
	(0.0057)	(0.0058)	(0.0056)	(0.0040)	(0.0041)	(0.0034)
Year	0.0100*		0.0105*	-0.0028*		-0.0008*
	(0.0004)		(0.0005)	(0.0003)		(0.0003)
Market Cap (in logs)	0.0936*	0.0942*	0.0850*	0.0689*	0.0695*	0.0543*
	(0.0034)	(0.0034)	(0.0039)	(0.0032)	(0.0032)	(0.0031)
Retail Share	0.6999*	0.7310*	0.6123*	0.2213*	0.2393*	0.1826*
	(0.0229)	(0.0238)	(0.0240)	(0.0149)	(0.0159)	(0.0126)
Mining, Quarrying, and Extraction	0.0078	0.0059		-0.0020	-0.0025	
	(0.0124)	(0.0124)		(0.0078)	(0.0078)	
Utilities	0.0140	0.0080		0.0356*	0.0323*	
	(0.0115)	(0.0114)		(0.0078)	(0.0078)	
Manufacturing	0.0028	0.0010		0.0095*	0.0089	
	(0.0076)	(0.0076)		(0.0048)	(0.0048)	
Retail	-0.0056	-0.0076		0.0023	0.0012	
	(0.0145)	(0.0144)		(0.0088)	(0.0087)	
Information	0.0151	0.0112		0.0355*	0.0333*	
	(0.0138)	(0.0134)		(0.0150)	(0.0147)	
Finance and Insurance	-0.0188	-0.0220*		-0.0187*	-0.0201*	
	(0.0108)	(0.0108)		(0.0064)	(0.0064)	
Other Sectors	0.0000	0.0000		0.0000	0.0000	
	(.)	(.)		(.)	(.)	
Year FE		✓			✓	
Firm FE			✓			✓
N	31359857	31359857	31359857	31359857	31359857	31359857