European Review of Agricultural Economics pp. 1–27
doi:10.1093/erae/jby026

The relationship between price and retail concentration: evidence from the US food industry

Vardges Hovhannisyan[†,*], Clare Cho[‡] and Marin Bozic[§]

[†]University of Wyoming, USA; [‡]United States Department of Agriculture, Washington, DC, USA; [§]University of Minnesota—Twin Cities, Saint Paul, MN, USA

Received June 2017; editorial decision May 2018; final version accepted June 2018

Review coordinated by Jack Peerlings

Abstract

This study utilises the product barcode, store and retail real estate data to obtain consistent estimates of the effects of retail market concentration on food prices in the USA. Our disaggregated data allow for an identification strategy that corrects for the endogeneity of concentration in the concentration–price relationship. Findings from an instrumental variables fixed-effects model indicate that prices rise with retail concentration, and that ignoring endogeneity results in a severe downward bias. A simulation analysis finds that a 5 per cent increase in concentration would increase prices by 18 per cent and decrease food consumption by 2–5 per cent. Our findings suggest mergers in the food industry could inadvertently lead to adverse effects.

Keywords: retail concentration, retail food price, endogeneity of retail concentration, instrumental variables fixed-effects regression

JEL classification: L11

1. Introduction

The global agri-food system has been undergoing significant structural changes recently. One such change has been the increasing consolidation and coordination at the retail level, which has resulted in the top 15 global supermarket companies generating over 30 per cent of the world's supermarket sales (USDA ERS, 2017a). Rising retail concentration has been particularly pronounced in the OECD countries, especially in the EU and the USA (Bukeviciute, Dierx and Ilkovitz, 2009; Balagtas, 2010). In the USA, traditional food retailers have been facing increased competition from nontraditional retailers, such as mass merchandisers and drug stores. The share of food sales from these nontraditional retailers increased from 13.7 to 21.5

*Corresponding author: E-mail: vhovhann@uwyo.edu

per cent from 2000 to 2011, while nonstore sales, such as home deliveries and direct sales from farms, remained consistently near 9 per cent during this period (USDA ERS, 2017b). Partially because of the competition from non-traditional retailers, large food retailers chose to save overhead costs by acquiring other retailers rather than creating new stores, resulting in a sharp increase in acquisitions since 1996 (Harris *et al.*, 2002). The consolidation trends in the 1990s intensified in the 2000s as the share of sales from non-traditional retailers increased, particularly among mass merchandisers (Wood, 2013). As a result, the combined market share among the four largest grocery chains in the USA increased from 16 to 32 per cent over the period 1982–2005; the remarkable growth continued thereafter, reaching nearly 40 per cent of the total market share in 2013 (Hovhannisyan, Stiegert and Bozic, 2014; USDA ERS, 2017b).

Similarly, the five largest food retailers in many EU member states were responsible for over 50 per cent of the retail sales in 2007, with even higher concentration levels among older member states (Bukeviciute, Dierx and Ilkovitz, 2009). Furthermore, rising concentration in certain EU states spilled over into neighbouring countries while many multinational retailers extended their reach internationally (Dobson, Waterson and Davies, 2003). This has raised the concern of negative effects for both suppliers and consumers, as retailers negotiate lower buying prices (particularly by offering their own private labels) and raise retail prices for consumers (Consumers International, 2012).

Retailers in many OECD countries have developed powerful private labels; expanding their services in prepared foods, in-store bakeries, floral departments and deli; offering services related to health and wellness, such as meal planning and cooking classes, and even creating contemporary interior store designs to provide a more appealing shopping environment (Rosenblum, 2014). The rise of powerful store brands, in particular, has provided successful large retail chains with unique competitive advantages over rival chains and national brand manufacturers (Steiner, 2004).[1] When combined with this non-price aspect of retail competition, the steady rise in concentration has the potential to reshape the horizontal competitive landscape among food retailers, as well as the vertical relationships between retailers and their suppliers. Although this may be beneficial for consumers, particularly if retailers grow larger and are able to use economies of scale, the increase in concentration could also reduce competition, thereby raising prices. Hence, rising retail concentration has been at the centre of heated debates among economists, policymakers, lawyers and industry stakeholders alike. As a result, the European Parliament's Committee on Economic and Monetary Affairs and the US Departments of Justice have organised public workshops with the goal of providing policymakers with an improved understanding of market

1 For example, a recent survey reveals that store brand adoption rate in the USA has increased from 12 to 41% over the time span 1992–2007 (Ipsos, 2007).

conditions that determine farm and consumer prices (US DOJ, 2011; European Parliament, 2016).

Despite a large body of literature on the relationship between market structure and firm performance, there is a paucity of evidence on the price effects of increasing concentration in the food retail industry (Richards and Pofahl, 2010; Hovhannisyan, 2017). Early studies investigate the effects of market concentration on firm profitability using industry-level cross-section data from a wide range of industries. The major finding emerging from this literature is that market concentration and firm market shares are positively related to firm profitability (Schmalensee, 1989; Singh and Zhu, 2008). Recent studies have shifted the focus from cross-industry to single-industry price–concentration analyses to sidestep issues that may stem from fundamental industry differences, profit measurement and the 'efficiency' critique put forth by Demsetz (1973) regarding firm superiority driving the positive relationship between market concentration and profit. Results from these price–concentration regressions indicate that prices go hand-in-hand with rising retail concentration, which is robust to the inclusion of economies of scale and organisational form (e.g. Lamm, 1982; Cotterill, 1986; Kaufman and Handy, 1989). More recent studies examining the USA (e.g. Cotterill, 1999; Hovhannisyan and Bozic, 2013, 2016) and the EU markets (e.g. Aalto-Setälä, 2002; Biscourp, Boutin and Vergé, 2013) further reinforce this finding.

An important challenge that remains largely unaddressed in studies of retail markets, where non-price dimensions play a significant role from competition and pricing perspectives, is the interpretation of the concentration effects on prices (Newmark, 2004). Therefore, most recent studies confine price–concentration analyses to food items that are sufficiently homogenous across retailers such as water, milk, flour and sugar. Biscourp, Boutin and Vergé (2013) present one such study based on panel data of retail food and non-food item prices in France. The panel nature of the data allows them to control for certain aspects of non-price competition that are considerably stable over time such as store location, amenities, quality of management and network effects.

Another important issue plaguing this literature and still presenting major challenges for estimation is endogeneity of market concentration due to the simultaneous nature in which prices and concentration are determined (e.g. Froeb and Werden, 1991; Berry, 1992). Unless accounted for, endogeneity can severely bias the estimate of the concentration parameter (e.g. Evans, Froeb and Werden, 1993; Manuszak and Moul, 2008). Given the lack of data on instruments that are both relevant and valid, the endogeneity of concentration remains largely unaddressed, although there have been attempts to alleviate the issue. For example, Biscourp, Boutin and Vergé (2013) construct an alternative measure of the Herfindahl–Hirschman index (HHI) for market concentration based on store size rather than revenue, which is generally less responsive to price incentives. With panel data becoming increasingly available, lagged HHI values have also become a popular choice for concentration instruments. As a case in point, Evans, Froeb and Werden (1993) apply a fixed-effects instrumental variables (IV) technique to examine the effects of rising concentration in the US airline industry. Similarly, Dafny, Duggan and Ramanarayanan (2012) rely on lagged HHI estimates to study the impact of increasing hospital concentration on insurance premiums. Some of the more recent studies employ a two-stage approach, where the first stage is used to derive a correction term for the endogeneity of market structure, which is then incorporated into the price–concentration regressions in the second stage (e.g. Manuszak and Moul, 2008; Zhu, Singh and Manuszak, 2009).

The current study aims at informing the discussion on Price—concentration in the US food retail industry while addressing a number of major shortcomings plaguing the previous literature. First, it adopts a combination of fixed-effects and IV techniques that account for store-level unobserved heterogeneity. This unobserved heterogeneity reflects time-invariant unobserved store characteristics, such as quality of service and management, location, amenities, transition zone (i.e. displays and other décor placed in front of stores) etc., which are important considerations when setting retail prices (Biscourp, Boutin and Vergé, 2013). Second, our empirical analysis is performed by combining a dataset with detailed store and product barcode-level price information with another dataset containing store characteristics of an exhaustive list of food retailers. Further, HHI estimates are computed using Nielsen TDLinx store characteristics data that cover an exhaustive list of food retailers. Third and, most importantly, our approach accounts for the endogeneity of retail concentration by exploiting unique retail real estate data on newly constructed retail space provided by Marcus & Millichap.[2] New construction is a particularly attractive instrument for concentration because it affects a retailer's fixed costs, whether it is a new retailer entering the market or an incumbent retailer renewing the lease. A final virtue of this study is that our identification strategy is easily applicable to the analyses of competition and prices in other countries with similar retail food systems, particularly the EU member states, and offers the promise of increasing the reliability of empirical findings.

2. A panel data model for food price and market structure

2.1. A reduced-form specification of price–concentration regression

We empirically investigate the relationship between retail food prices and market structure by adopting a combination of IV and fixed-effects econometric techniques. Our base specification follows the reduced-form framework suggested by Biscourp, Boutin and Vergé (2013), where retail prices are expressed as a function of HHI as a measure of market concentration and market-specific descriptors such as population and income:

2 Marcus & Millichap is a commercial real estate brokerage firm and one of the largest US companies specialising in real estate investment services. It also conducts research on commercial real estate markets. See Section 3 for additional details.

$$p_{i,j}^{y,m} = \lambda^y X_{c(j)}^y + \delta^y \mathrm{HHI}_{c(j)}^y + \alpha_j + \psi_i^{y,m} + \gamma_{\varphi(j)}^y + \varepsilon_{i,j}^{y,m} \tag{1}$$

where $p_{i,j}^{y,m}$, price (in logarithm) of product i in store j in month m of year y; $X_{c(j)}^y$, vector of market characteristics (other than concentration such as population and income) relating to market c where store j is located in year y; $\mathrm{HHI}_{c(j)}^y$, HHI of market concentration; α_j, store fixed effects; $\psi_i^{y,m}$, dummy variables representing interactions between product, year and month; $\gamma_{\varphi(j)}^y$, store-type effects with $\varphi(j)$ denoting the retail format to which store j belongs (i.e. convenience store, mass merchandiser, discount store etc.) and $\varepsilon_{i,j}^{y,m}$, *i.i.d.* disturbance for product i in store j in month m of year y.

Following Biscourp, Boutin and Vergé (2013), we estimate the cross-section specification using interaction dummies for product, year and month as additional covariates. We also include store fixed effects, which account for unobserved store heterogeneity. Our goal with this specification is to control for the effect of store-specific factors, such as amenities, location and quality of staff, as well as market-level unobserved factors that remain unchanged over time. Finally, all the specifications also account for store-format effects and interaction dummies for product and store-format (allowing for the possibility of certain products being priced differently based on the store-format) are used for robustness checks (Biscourp, Boutin and Vergé, 2013).

Although the base specification of our reduced-form model is structured following Biscourp, Boutin and Vergé (2013), the implementation of our models is different. For example, the store types identified in Biscourp, Boutin and Vergé (2013) differ from those available in our dataset (e.g. hard discount vs. dollar store). In addition, Biscourp, Boutin and Vergé include an indicator variable for store brand products while our paper examines the average price across barcodes for all products that can be considered sufficiently homogenous across store types, such as water or milk, as discussed further in Section 3. Finally, we explicitly account for the endogeneity of market concentration by utilising the portion of variation in the concentration driven by an exogenous supply shifter, namely newly constructed retail space.

2.2. Endogeneity of retail concentration and identification strategy

Endogeneity of market concentration may arise through the correlation between concentration ($\mathrm{HHI}_{c(j)}^y$) and the *i.i.d.* disturbance ($\varepsilon_{i,j}^{y,m}$).[3] As a firm becomes more efficient or gains more market power, the overall market becomes more concentrated and firm performance (e.g. profitability, price levels) is affected as well. It could also be performance feeding back into structure through its impact on firm conduct (e.g. various promotional campaigns, decisions regarding entry, exit and investment in new capacity etc.), causing the industry structure to evolve over time (Evans, Froeb and Werden,

3 Fixed-effects estimation addresses the correlation between concentration and store-specific variables.

1993). Further, revenue and output-based HHI are calculated using firm output and revenue, which are determined simultaneously with prices.[4] This results in biased and inconsistent OLS estimates. The consensus appears to be that the OLS estimator underestimates the association between price and concentration (e.g. Manuszak and Moul, 2008). In general, nevertheless, the direction of the bias cannot be predicted and the effect largely depends on various modelling assumptions, as well as certain market characteristics (Froeb and Werden, 1991).

This study combines fixed-effects and IV procedures to address the endogeneity of retail concentration. A major challenge that needs to be addressed is the choice of an instrument. Ideally, a valid concentration instrument satisfies the requirements of: (i) relevance, when there is sufficient correlation between a measure of concentration and the instrument and (ii) exogeneity, when the instrument is properly excluded from the reduced-form price equation, thus being uncorrelated with unobserved price determinants and affecting prices indirectly through the instrument's impact on concentration. However, the common practice in the previous literature has been relying on imperfect instruments such as lagged HHI values, mostly due to limited data (e.g. Evans, Froeb and Werden, 1993; Dafny, Duggan and Ramanarayanan, 2012). Our approach, on the other hand, provides a correction for the endogeneity of retail concentration by exploiting the variation in retail completion or newly constructed retail space using unique retail real estate data. This variable is a particularly attractive instrument for the retail concentration because of the nature of its relationship with a retail fixed cost – retail rent.

Retail rents constitute fixed payments that remain constant in the duration of rental contracts and account for a sizable portion (8.3 per cent) of total operating costs of US food retailers (Annual Retail Trade Survey, U.S. Census Bureau, 2012). As such, retail rent is a significant a factor in a new firm's decision to enter the market or an incumbent firm's decision to renew the lease, which invariably affects market concentration, *ceteris paribus* (Newman and Cullen, 2002). In turn, retail rents are determined as an equilibrium outcome in the retail rental markets by the interaction of *retail space supply* and demand factors.

Our identification strategy exploits the supply-side variation in retail rents that is driven by new retail space construction. New construction refers to completions, or the total square footage in all new buildings that have passed the final inspection under the building permit during the period under consideration and are ready-to-use (Mourouzi-Sivitanidou, 2002). Given the long life of real estate assets, new construction and depreciation have been discovered to be the two most important factors determining how real estate markets and the retail real estate inventory evolve over time (Gyourko and Voith, 1993; Rmourouzi-Sivitanidou, 2011). Therefore, *new construction* is a key supply-side factor in real estate market analyses and a key determinant of

4 Biscourp, Boutin and Vergé (2013) use an alternative measure of market concentration, i.e. capacity-based HHI, which sidesteps this particular source of endogeneity.

retail rents, and should satisfy the relevance requirement for the concentration instrument.[5]

The exogeneity of new construction relies primarily on the institutional characteristics of the retail real estate market. Specifically, empirical evidence emerging from the real estate economics research identifies land availability, development costs, tax law changes, capital market cycles, interest rates as well as demographic, sociological and local trends as some of the most important factors affecting retail space supply through their impact on capital costs and supply-side constraints (see, for example, Segal and Srinivasan, 1985; Benjamin *et al.*, 1998). When combined with informational inefficiencies, investment-decision lags and inherently long construction lags, these supply-side constraints make retail space supply highly inelastic to the changing economic climate and irresponsive to changing levels of retail sales (Benjamin, Jud and Okoruwa, 1993; Benjamin, Jud and Winkler, 1995; Eppi and Shilling, 1996). This process can be further prolonged by government policies regarding building permits, growth regulations, local zoning and other institutional controls that may hamper retail real estate development. As a result, it usually takes about 2 years to complete construction from the time of inception and 1–2 years for lease-up (Sivitanidou and Sivitanides, 2000). These features taken together minimise the possibility of food prices feeding back into retail construction, which is essential from the identification perspective.

3. Data description

We compile data on retail food prices, store and market characteristics, as well as newly constructed retail space from a variety of sources: (i) IRI scanner data that contain store and product barcode-level information on retail dollar sales and quantity for food and beverage products marketed throughout the USA;[6] (ii) Nielsen TDLinx, which provides annual store characteristics data for an exhaustive list of food retailers; (iii) Bureau of Economic Analysis of the US Department of Commerce that provides annual market characteristics data regarding population and per capita income and (iv) Marcus & Millichap real estate data on newly constructed retail space. We provide a more detailed description of these datasets in the Appendix.

We base our empirical analysis on nearly 2,000 food products in the IRI that are sufficiently homogenous across stores and dates (e.g. water, milk, sugar, salt etc.) to sidestep potential issues related to the effects of product differentiation on food prices (Biscourp, Boutin and Vergé, 2013). To ensure empirical tractability,

5 We recognise the possibility of rents also containing a variable component (i.e. a percentage of retail sales volume), in addition to a *fixed* base payment (Benjamin and Chinloy, 2004). While the variable component may be responsive to food prices and sales volumes, our identification strategy centres on this *fixed component* that is exogenous to retail food price determination in the short-run.

6 Some of the retailers in IRI only provide sales information on a regional marketing area level. We only included those retailers with store-level price data.

Table 1. Descriptive statistics for the retail newly completed area by market, 2008–2012

Market	State	New space (footage^2)			New space/total space (%)		
		Mean	SD	CV (%)	Mean	SD	CV (%)
Charlotte	NC	827	592	71.58	12.26	7.90	64.46
Chicago	IL	2,102	1,934	92.01	15.98	13.27	83.05
Cincinnati	OH	793	823	103.78	10.88	9.90	91.03
Columbus	OH	448	257	57.37	6.51	3.40	52.29
Dallas	TX	3,300	2,923	88.58	39.41	31.83	80.78
Houston	TX	2,103	1,926	91.58	9.92	8.33	83.95
Indianapolis	IN	630	661	104.92	7.94	7.74	97.47
Jacksonville	FL	615	758	123.25	8.08	8.97	110.93
Louisville	KY	158	98	62.03	2.45	1.35	54.92
Milwaukee	WI	660	273	41.36	10.82	3.78	34.94
Minneapolis	MN	596	321	53.86	7.23	3.74	51.80
New York	NY	1,079	450	41.71	15.34	5.95	38.82
Phoenix	AZ	2,349	2,907	123.75	24.88	27.29	109.69
Sacramento	CA	715	467	65.31	12.92	7.74	59.90
San Antonio	TX	1,375	1,458	106.04	13.66	13.44	98.38
San Diego	CA	301	114	37.87	3.72	1.24	33.42

Source: Marcus & Millichap Real Estate Data, 2008–2012.
Note: Retail real estate data vary annually over the period 2008–2012.

we aggregate these food items into 129 food groups (see Online Appendix, in supplementary data at *ERAE* online). Our final dataset contains monthly prices over the period 2008–2012, which results in unbalanced panel data that cover the following retail store formats: convenience stores, dollar stores, drug stores, grocery stores and mass merchandisers.[7] Details on the sales composition by store type are provided in Table A1.

Given the large number of food products (129) and retail stores (3,144) in our sample, we confine our empirical analysis to 16 US metropolitan statistical areas (Table 1). For identification purposes, we select retail markets with the highest annual variability of the number of retailers (i.e. retailer entry and exit) in the sample period. As can be observed from Table A2, markets vary widely in terms of market concentration; however, most markets appear to be low to moderately concentrated. The underlying reason may be the way the markets are defined. Further, spatial variation seems to outweigh the temporal variation with the coefficient of variation ranging from as low as 2.25 per cent for Minneapolis, MN to 10.35 per cent for San Diego, CA.

As shown in Figure 1, prices for coffee, water and eggs seem to be positively correlated with retail concentration, while milk prices and concentration do not appear to be following a particular pattern. This may happen, for example, when milk pricing follows loss leader pricing with the aim of

7 Descriptions of these store types are provided in the Appendix.

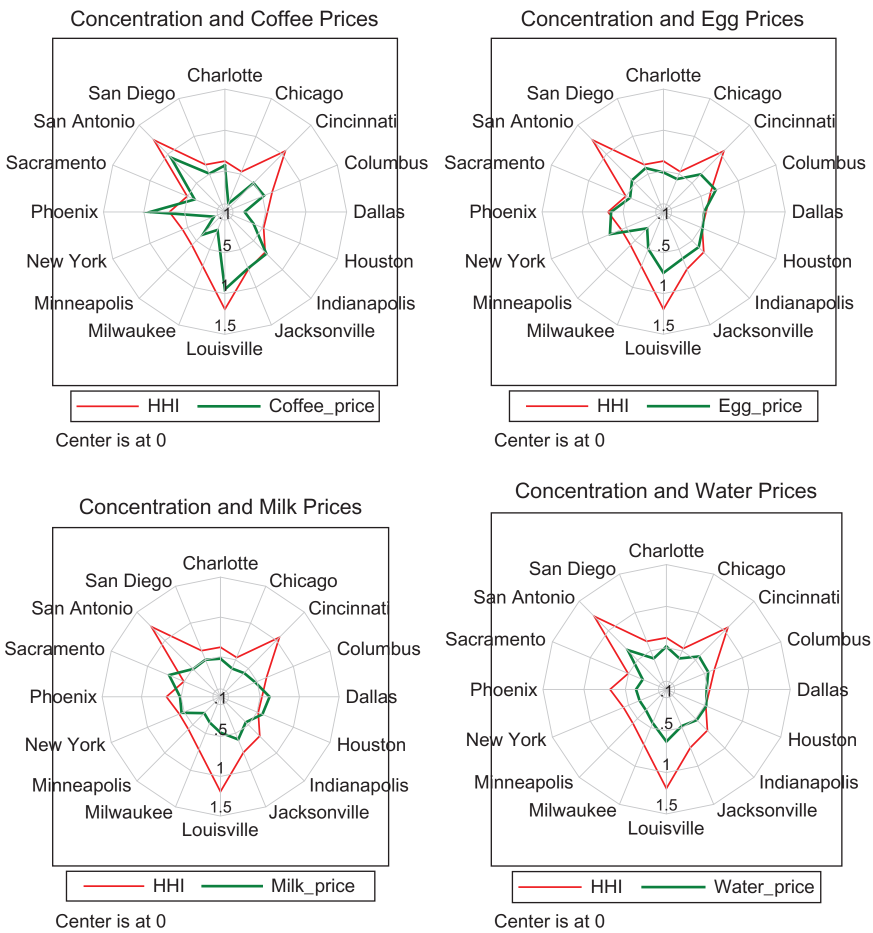

Fig. 1. Market concentration (HHI) and selected prices by the market.
Note: Both HHI and prices are expressed as market averages over 2008–2012.

attracting a certain level of store traffic (Green and Park, 1998). While this graphical analysis can shed light on how concentration and prices may be interrelated, nevertheless, it omits the effects of certain important variables (e.g. retail costs, regulations) and thus is not an accurate representation of concentration–price relationship.

We exploit real estate data provided in the annual *National Retail Reports* by Marcus & Millichap to construct instruments for market concentration. More specifically, we utilise newly constructed retail space measured in square feet for the 16 US retail markets in our sample for the period 2008–2012. The completion of new retail space manifests considerable spatial and temporal variability during our sample period with the coefficient of variation ranging from 41.36 per cent for Milwaukee, WI to 123.75 per cent for Phoenix, AZ (Table 1, left panel). To put the instruments for concentration in a more meaningful context, we construct the ratio of new retail space completion for a given market and year to the respective total retail space available (Table 1, right panel).[8] Sampling variability of this ratio is also noteworthy with the mean share of the new retail space varying from 2.45 per cent for Louisville, KY to 39.41 per cent for Dallas, TX. Further, the respective coefficient of variation extends from 33.42 per cent for San Diego, CA to 110.93 per cent for Jacksonville, FL. It deserves noting that, in our econometric specification, the identification of the effect of concentration on prices relies heavily on the exogenous variation of these instruments.

4. Empirical results and policy implications

In this section, we provide the results from concentration–price regressions. We present the results from a simple OLS regression and a more sophisticated fixed-effects estimation. Further, we estimate both with an IV. Finally, we briefly discuss the major policy implications stemming from our findings.

4.1. Empirical relationship between retail price and retail concentration

Our empirical results are provided in Tables 2 and A3. For benchmark comparison, we first report the estimates from the cross-section/OLS regressions, which disregard both unobserved store heterogeneity and the endogeneity of concentration (Table A3, left panel). These results confirm that mass merchandisers offer the lowest prices on a wide spectrum of food products among the retail formats in our sample (with the exception of grocery stores in 2008). Moreover, the price gap between mass merchandisers and the remaining formats, for the most part, has been on a rise over the 5-year span. These findings are also supported by the IV-OLS results (Table A3, right panel). Based on the estimated coefficients for population, a 1 per cent increase in the population is associated with 0.254–0.276 per cent increase in food prices under OLS, while the price change magnitude falls in 0.659–0.745 per cent under IV-OLS. Similarly, food prices rise by 0.017–0.057 per cent with a 1 per cent increase in consumer income under OLS and by 0.121–0.225 per cent under IV-OLS. Finally, market concentration and food prices do not appear to have a clear relationship under OLS. By contrast, under IV-OLS, prices are found to go hand-in-hand with market concentration with a one standard deviation increase in HHI causing food prices to go up by 0.032–0.332 per cent.[9] However, IV-OLS still suffers from omitted variable

8 We calculate total retail space from the respective markets based on the Nielsen TDLinx data on store characteristics.

9 To allow for a more meaningful interpretation of the concentration effects, we follow Cotterill (1999) to standardise the HHI variable (i.e. the HHI coefficients are now interpreted as reflecting the percentage change in prices for a one standard deviation change in HHI).

Table 2. Parameter estimates and standard errors from fixed-effects and fixed-effects IV regressions

	Fixed-effects					IV fixed-effects				
	2008	2009	2010	2011	2012	2008	2009	2010	2011	2012
Convenience	Ref.	**0.1839**	**0.3373**	**0.5023**	**0.5918**	Ref.	**0.2039**	**0.3610**	**0.5499**	**0.5959**
		(0.0103)	*(0.0167)*	*(0.0261)*	*(0.0313)*		*(0.0126)*	*(0.0192)*	*(0.0359)*	*(0.0311)*
Dollar	Ref.	**0.0573**	**0.1754**	**0.3011**	**0.2942**	Ref.	**0.0620**	**0.1813**	**0.2963**	**0.3043**
		(0.0096)	*(0.0167)*	*(0.0265)*	*(0.0318)*		*(0.0114)*	*(0.0187)*	*(0.0331)*	*(0.0324)*
Drug	Ref.	**0.1010**	**0.1987**	**0.4227**	**0.5496**	Ref.	**0.1017**	**0.2025**	**0.4536**	**0.5428**
		(0.0095)	*(0.0167)*	*(0.0263)*	*(0.0316)*		*(0.0111)*	*(0.0183)*	*(0.0345)*	*(0.0312)*
Grocery	Ref.	**0.0337**	**0.1535**	**0.3713**	**0.4455**	Ref.	0.0140	**0.1333**	**0.2998**	**0.4595**
		(0.0096)	*(0.0169)*	*(0.0269)*	*(0.0325)*		*(0.0118)*	*(0.0196)*	*(0.0329)*	*(0.0354)*
Mass merchandisers	Ref.	Ref.	Ref.	Ref.	Ref.	Ref.	Ref.	Ref.	Ref.	Ref.
HHI	–0.0221	**–0.0364**	–0.0198	–0.0120	–0.0176	**0.2393**	**0.1733**	**0.1791**	**0.0351**	**0.2817**
	(0.0150)	*(0.0153)*	*(0.0154)*	*(0.0151)*	*(0.0154)*	*(0.0731)*	*(0.0760)*	*(0.0760)*	*(0.0091)*	*(0.0973)*
Population	0.2105	0.2113	**0.2322**	**0.2299**	**0.2350**	**0.5465**	**0.5448**	**0.5634**	**0.4923**	**0.5657**
	(0.1109)	*(0.1115)*	*(0.1113)*	*(0.1111)*	*(0.1107)*	*(0.2229)*	*(0.2261)*	*(0.2265)*	*(0.2130)*	*(0.2228)*
Income	**0.0367**	**0.0222**	0.0236	**0.0226**	0.0007	**0.2534**	0.1894	**0.1833**	**0.1420**	**0.1837**
	(0.0087)	*(0.0080)*	*(0.0770)*	*(0.0080)*	*(0.0088)*	*(0.0401)*	*(0.3230)*	*(0.0307)*	*(0.0235)*	*(0.0366)*
Additional controls	(Year × month, product × year, store fixed-effects)					(Year × month, product × year, store fixed-effects)				
Number of observations	1,457,861	1,525,662	1,542,870	1,603,345	1,625,428	1,457,861	1,525,662	1,542,870	1,603,345	1,625,428
First-stage *F*-statistics[a]	*NA*					113.3				

Notes: Values in bold identify parameter estimates statistically different from 0 at the 0.05 level of significance. Store-level cluster-robust standard errors are italicised. The final sample includes 7,755,166 observations.

[a]Presents the Stock–Wright *F*-statistic of zero coefficient on new retail space construction.

bias through unobserved store heterogeneity, which can confound the estimated effects of concentration on prices.

The recent literature provides ample evidence pointing to the significant role that unobserved store heterogeneity plays in retail price formation (e.g. Evans, Froeb and Werden, 1993; Biscourp, Boutin and Vergé, 2013; Lin and Wang, 2015). For this reason, we further utilise a fixed-effects estimator (i.e. within estimator) in our analysis, which allows us to control for unobserved store characteristics that are constant or likely to remain stable over time (e.g. store location, size, amenities, distribution effects, quality of management etc.). Based on a Hausman test, we may reject the null hypothesis that all of the coefficients from the fixed-effects and random-effects panel data models are the same (the value of the associated test statistic is $\chi^2(3, 245) = 13{,}441$, or equivalently, p-value $= 0.00$).[10] This implies that $\mathrm{Cov}(\alpha_j,\ \mathrm{HHI}^y_{c(j)}) \neq 0$ and that the OLS and random-effects models yield inconsistent estimates. Therefore, the fixed-effects regression is our preferred specification.

Estimation results from the fixed-effects regressions are provided in Table 2. Our inclusion of store fixed-effects requires that the store-format effect on prices be dropped for one of the years, i.e. 2008 in this case, to avoid the dummy variable trap. This implies that the coefficients capturing particular store-format effects are now interpreted as reflecting the difference between the prices at this retail format for that year on the one hand and those at mass merchandisers (i.e. reference format) for the entire sample period along with the remaining formats in 2008 on the other. Excluding a dummy variable from the time effects (i.e. the reference year) and the store fixed-effects (i.e. the reference store format) allows for a more meaningful comparison between the prices at different formats.

As has been found in the previous literature and further illustrated above, market concentration remains endogenous to price determination even after accounting for unobserved store heterogeneity. Hence, we also run our regression using an IV fixed-effects model based on the concentration instruments constructed in this study.[11] Using a Durbin–Wu–Hausman test procedure, we find strong empirical evidence that concentration is endogenous to

10 We acknowledge that the Hausman test may be neither a necessary, nor sufficient statistic for deciding between fixed and random-effects estimators, as illustrated by Clark and Linzer (2015) based on a series of simulation experiments.

11 Assuming exogenous market concentration, the estimated coefficients associated with population (0.230–0.235) and per capita income (0.022–0.037) suggest that both are positively associated with price in the fixed-effects model (Table 2, left panel), although coefficients for certain years are found to be insignificant. Importantly, a majority of the HHI estimated coefficients are statistically insignificant, with the exception of that for 2009, which is negative and insignificant (−0.036). This is in contrast to other similar studies, such as Froeb and Werden (1991) and Evans, Froeb and Werden (1993), and may be reflective of the fact that unobserved variation in factor prices exceeds the unobserved variation in demand in the OLS specification. This has been illustrated by Froeb and Werden (1991), with the net bias being positive (i.e. −0.036 in OLS model vs. −0.009 in fixed-effects model).

price formation.[12] Further, the first stage estimation results reveal that new retail space construction is strongly correlated with market concentration from 2008 to 2012 under both the cross-section and the fixed-effects models with the respective F-statistics values of 122.4 and 113.3 (Tables 2 and A3). This finding is robust to multiple tests under the fixed-effects model, including the Andersen–Rubin Wald test ($\chi^2(10) = 90.8$), Stock–Wright LM test ($\chi^2(10) = 113.3$), Cragg–Donald Wald test ($\chi^2(10) = 9{,}174.6$) and Kleibergen–Paap Wald *rk* test that yields heteroscedasticity-robust results ($\chi^2(10) = 59.1$).[13] Therefore, our instruments satisfy the relevance requirement, and we dismiss concerns about weak instrument bias (Stock and Yogo, 2005).

Estimation results from the IV fixed-effects model with the associated store-level cluster-robust standard errors are presented in Table 2, right panel.[14] It can be observed that a great majority of the coefficients of the retail formats trace the fixed-effects estimates closely (Table 2, left panel). HHI estimates present an important exception in that they exceed in magnitude the respective estimates from the fixed-effects model and are all positive and statistically significant for the entire sample period. Specifically, the estimates from the IV model range from 0.1733 to 0.2817 (e.g. the fixed-effects coefficient for year 2012 is about one/tenth of the respective IV fixed-effects coefficient). An interesting finding that emerges is that the HHI coefficients from the IV fixed-effects regression manifest a decline over the period 2008–2009 from 0.2393 to 0.1733, which is followed by a steady increase through 2012. This might be indicative of food demand becoming less inelastic in the aftermath of the 2008 recession, thus intensifying retail competition in the short-run (based on a joint test for equality of parameters ($F(4, 3{,}143) = 9.24$, p-value $= 0.00$), we may reject the null hypothesis of equality).

As a robustness check, we also estimate the fixed-effects and IV fixed-effects specifications that include three high-dimensional fixed effects. Specifically, following Biscourp, Boutin and Vergé (2013), we include interaction terms between products and retail formats in addition to the store fixed-effects and the interaction of year and month variables.[15] This allows

12 The Durbin–Wu–Hausman test determines whether the estimates from the fixed-effects and the IV fixed-effects models are statistically significantly different. The coefficient for the residuals in the second stage is statistically significant, implying that we may reject the null that concentration is exogenous. However, it is important to note that this result is conditional on the validity of our instrument.

13 We do not perform a test for the overidentifying restrictions since our equation is exactly identified. It is worth noting that overidentifying restrictions provide little information on the validity of instruments (Parente and Silva, 2012). Instead, the validity of instruments should be based on the underlying economic model and the causal mechanism such as the one presented earlier. Further, overidentifying restriction tends to be rejected in empirical applications with considerable parameter heterogeneity.

14 As illustrated by Angrist and Pischke (2009), the finite sample bias of the formula for homoscedastic errors is smaller than that of the robust sandwich estimator. Nonetheless, our very large sample size suggests using the robust estimator.

15 We estimate this multiple high-dimensional fixed-effects model using the REGHDFE Stata programme developed by Correia (2017), which is based on memory-saving techniques and requires considerably less run-time vis-à-vis the standard panel data estimation programmes.

for the possibility of food products being priced differently across the retail formats (Biscourp, Boutin and Vergé, 2013). Overall, the parameter estimates from this more general model are considerably close to those reported in Table 2 in terms of the sign and magnitude.[16]

We further explore the possibility of price effects of concentration varying by format. Specifically, we interact HHI with store-format dummies to capture the differential effects of format on price. Our results show that prices decline with rising concentration of convenience stores (−0.1019 to −0.0649) and drug stores, while dollar store (0.0492–0.0626) and grocery store (0.0758–0.1521) concentration is positively related to prices (Table A4). This may be due to a relatively smaller size of convenience and drug stores; increasing store size up to a certain threshold leads to economies of scale and cost savings that are passed on to consumers. In contrast, grocery retailing is an arena of constant power struggle where power lies increasingly with big companies. With fewer and even larger retail companies, therefore, our findings may be indicative of increasing pricing power on the part of grocery stores, which translates into the exercise of market power. The results are eye opening in that rising HHI may have unequivocal impact on prices with the direction of effects largely depending on the source of the increase in HHI. It deserves noting that proper control of the endogeneity of format-specific concentration is more data demanding and presents considerable challenges. Specifically, given that concentration instrument is measured at a more aggregate level (i.e. market-level) vis-a-vis format-specific concentration, it can affect prices through more than one format and, thus, can confound the concentration effects. A proper specification would examine the effects of format-specific concentration on price, where instruments reflect exogenous variation in concentration and are format specific.

Finally, to get a better sense of the implications of rising retail concentration for consumer welfare, we perform a simple simulation analysis that is inspired by Hernandez and Torero (2013). Specifically, we explore the potential effects of a concentration-induced increase in food prices on consumer food consumption. Assuming two hypothetical scenarios of concentration rise, we predict the implied price change based on δ^y estimates, which are then used in conjunction with price elasticity and predicted food consumption response estimates provided by Okrent and Alston (2012) to quantify the effects of concentration on food consumption.[17] The specific scenarios considered include an increase in retail concentration by one and two standard deviations, which, based on the predictions of our model, will raise food prices by 18.17 and 36.34 per cent, respectively (i.e. one standard deviation is equivalent to the 5 per cent of the average HHI; Table 3). Based on unconditional price elasticities and the resulting food consumption predictions in

16 The results from this specification are available upon request.

17 Our approach is similar to Hernandez and Torero (2013) in that it relies on price elasticity and predicted consumption estimates obtained from a formal demand analysis.

Table 3. Projected impact of rising market concentration on food price and consumption

Variable	Scenario 1	Scenario 2
Hypothetical change in HHI (standard deviation)	1	2
Concentration-induced price change (%)	18.17	36.34
Apples		
Average price elasticity	–0.58	–0.58
Concentration-induced change in consumption (%)	–3.63	–7.26
Citrus		
Average price elasticity	–1.10	–1.10
Concentration-induced change in consumption (%)	–4.72	–9.44
Processed fruits and vegetables		
Average price elasticity	–0.77	–0.77
Concentration-induced change in consumption (%)	–1.66	–3.32
White bread		
Average price elasticity	–1.54	–1.54
Concentration-induced change in consumption (%)	–3.54	–7.08
Rice and pasta		
Average price elasticity	–0.07	–0.07
Concentration-induced change in consumption (%)	–4.59	–9.18
Fish		
Average price elasticity	–0.84	–0.84
Concentration-induced change in consumption (%)	–4.36	–8.72

Note: Concentration-induced price change is computed as the average of δ^y in equation (1) over 2008–2012. To calculate concentration-induced consumption change, we rely on unconditional price elasticities and predicted food consumption change (taking into account cross–price relationships between food products) estimated by Okrent and Alston (2012).

Okrent and Alston (2012) that take full account of own and cross–price effects, we simulate the impact of rising concentration on the consumption of several food products. For example, our predictions indicate a remarkable decline in apple consumption with a one standard deviation increase in HHI reducing apple consumption by 3.63 per cent. This change is even more pronounced for citrus consumption with 4.72 per cent, a 9.44 per cent expected changes associated with a one and two standard deviation changes in concentration. We further obtain predicted consumption changes for processed fruit and vegetables (−1.66 per cent per standard deviation), white bread (−3.54 per cent), rice and pasta (−4.59 per cent) and fish (−4.36 per cent).

Despite these important insights, we recognise that our simulation exercise is far from a complete policy analysis, for which structural models of retail market behaviour are better suited. Rather, it represents a simple demonstrative tool to provide important information about the likely effects of firm expansion and mergers in food retail industries. Further, we acknowledge that our simulation analysis does not take into consideration vertical interactions between retailers and upstream suppliers and the potential effects thereof on food prices. For example, these may be cost savings resulting from countervailing power of large retail chains when economies of scale and scope are used to negotiate lower prices to upstream suppliers, a share of which is subsequently transferred to final consumers. Finally, we acknowledge the danger associated with extrapolation of the HHI effects on food prices (i.e. predicting price response to HHI values that lie outside of the range used to fit the model). Thus, our results may overestimate the price effects of rising retail concentration and should be interpreted with caution.

4.2. Policy implications

Competition in food systems has always been an important public policy issue in a number of developed countries. In the USA, in particular, it has been at the centre of heated debates among agricultural producers, food processors, economists and lawmakers following important structural changes in the food marketing system. Retail competition has been of particular interest, given the recent steady rise in the retail concentration and consolidation and the important implications thereof for consumer welfare.

Our results indicate that retail food prices go hand-in-hand with retail concentration and that prices rise faster than they would appear under conventional research methods. By confining our focus to products that are largely homogenous, and by accounting for store-specific and temporal effects, our interpretation of this finding is that rising retail concentration is detrimental to retail competition and ultimately leads to higher food prices through the exercise of retail market power. In line with the previous literature, this finding suggests that the underlying factor driving mergers in the food retail industry is not the retailer motivation to utilise economies of scale, but rather to enhance firm market power. This key insight seems to be readily generalisable to many EU member states, notwithstanding country-specific constraints and idiosyncrasies, given the similar pattern of retail concentration in the EU and the USA. This hypothesis is further reinforced by a number of similar studies focusing on the EU market (see, for example, Aalto-Setälä (2002) for Finland, Biscourp, Boutin and Vergé (2013) for France and Ciapanna and Rondinelli (2014) for nine European countries). Therefore, more scrutiny should be exercised when evaluating potential mergers in food industries to avoid distorted policy outcomes and unfavourable effects on consumer welfare.

A final virtue of this study lies in its applicability to the evaluation of concentration effects in other countries. This is especially true for the EU countries where improving retail scanner and other data availability makes it possible to account for a variety of econometric issues including concentration endogeneity, thus contributing to improved decision-making in public policy arena.

5. Summary and conclusions

Rising retail concentration in a number of OECD countries, and especially in the EU and the USA, has the potential to reshape the competitive landscape in final goods. When an industry evolves towards greater concentration, it usually means that firms either seek market power that confers higher output prices on their buyers, or lower input prices on the procurement side, or the firms actively seek scale or scope economies to gain in efficiency. Therefore, the retail price–concentration relationship remains an empirical question, which we investigate in this application using the US food retail industry as our empirical setting.

The current study contributes to this discussion by offering a more credible identification of the effects of retail concentration on food prices. Our uniquely disaggregated data allow for the identification strategy that accounts for the type of endogeneity that plagues many previous studies on price–concentration relationship. Specifically, we employ a combination of fixed-effects and IV techniques that account for both store-level unobserved heterogeneity and the endogeneity of retail concentration. We rely on real estate economics literature to identify an exogenous supply-side variation in retail concentration, i.e. newly constructed retail space, and to establish its relationship to retailer size distribution and retail food prices through a causal mechanism.

Our findings indicate that rising retail concentration results in higher food prices, and that prices rise faster than conventional research methods suggest. This finding is more robust to the choice of a product handled by food retailers, given the focus of our study on over a 100 food groups spanning nearly 2,000 individual food items; previous studies offer considerably narrower product coverage, often limited to a single category. We interpret this finding as evidence of retailers seeking higher market power through mergers with rival firms and chains. Given the very similar retail industry dynamics in the EU and the USA, this key insight seems to be readily generalisable to many EU member states. Further, our qualitative results agree with a number of similar studies focusing on the EU market.

This study shows that it is important to consider potential adverse effects from mergers in the food industry. This is especially important for a considerable number of European countries, where high levels of concentrations have raised concern that food retailers are able to negotiate lower buying prices from suppliers and set higher retail prices for consumers (Consumers International, 2012). As demonstrated by workshops at both the European Parliament's Committee on Economic and Monetary Affairs and the US Departments of Justice, there may be various effects of increasing concentration in food retail that researchers can examine to inform policymakers of the effect on both consumers and suppliers (European Parliament, 2016; U.S. DOJ, 2011).

We acknowledge that the current study is not free from limitations. Specifically, we confine our analysis to the short-run effects of retail concentration (i.e. assuming instantaneous price reactions) to sidestep potential identification issues stemming from dynamic considerations. Further, in an attempt to reduce the computational burden of our empirical framework, we limit our sample to a 5-year period extending from 2008 to 2012, assuming away measurement error. However, when it is present, the relatively short temporal variation can magnify measurement error.

Future work would benefit substantially from a more complete analysis of concentration–price relationship that would take firm and industry dynamics into account either through lagged effects in a reduced-form framework or a full-blown structural model based on retailer behavioural assertions. Extending the sample period as more recent data become available and computational capacity continues to improve would further increase the value added by these studies. Future research should consider the differential effects of store-format specific concentration on food prices, where data on format-level concentration instruments are readily available. Finally, a proper specification for examining the effects of format-specific concentration on price would require the use of instruments that reflect exogenous variation in concentration and are format specific.

Supplementary data

Supplementary data are available at *European Review of Agricultural Economics* online.

Disclaimer

Any opinions, findings, recommendations or conclusions are those of the authors and do not necessarily reflect the views of the Economic Research Service, U.S. Department of Agriculture. The analysis, findings and conclusions expressed in this paper also should not be attributed to either Nielsen or Information Resources, Inc. (IRI). This research was conducted in collaboration with USDA under a Third Party Agreement with IRI.

Acknowledgements

Vardges Hovhannisyan is grateful to Dale Menkhaus, Alexandre Skiba, Chris Bastian, Alessandro Bonanno, as well as the seminar participants at the University of Wyoming, Colorado State University, and the RTI, AAEA, and USDA joint workshop. The authors would like to thank the Editor, Dr. Jack Peerlings and three anonymous referees for helpful comments. All remaining errors are our own.

References

Aalto-Setälä, V. (2002). The effect of concentration and market power on food prices evidence from Finland. *Journal of Retailing* 78: 207–216.

Angrist, J. and Pischke, J. S. (2009). *Mostly Harmless Econometrics*. Princeton, NJ: Princeton University Press.

Balagtas, J. V. (2010). Changing structure and competition in food and agricultural markets. *Choices (New York, N.Y.)* 25(2).

Barros, P. P., Brito, D. and de Lucena, D. (2006). Mergers in the food retailing sector: an empirical investigation. *European Economic Review* 50: 447–468.

Benjamin, J. and Chinloy, P. (2004). The structure of a retail lease. *Journal of Real Estate Research* 26(2): 223–236.

Benjamin, J.D., Jud, G.D. and Okoruwa, A.A. (1993). Forecasting the stock of retail space using the Koyck distributed lag model. *Journal of Property Research* 10(3): 185–192.

Benjamin, J.D., Jud, G.D. and Winkler, D.T. (1995). An analysis of shopping center investment. *The Journal of Real Estate Finance and Economics* 10(2): 161–168.

Benjamin, J., Jud, D. and Winkler, D. (1998). A simultaneous model and empirical test of the demand and supply of retail space. *Journal of Real Estate Research* 16(1): 1–14.

Berry, S. T. (1992). Estimation of a model of entry in the airline industry. *Econometrica: Journal of the Econometric Society* 60(4): 889–917.

Biscourp, P., Boutin, X. and Vergé, T. (2013). The effects of retail regulations on prices: evidence from the Loi Galland. *The Economic Journal* 123: 1279–1312.

Bukeviciute, L., Dierx, A. and Ilkovitz, F. (2009). The functioning of the food supply chain and its effect on food prices in the European Union. European Economy, Occasional Papers No. 47. European Commission.

Ciapanna, E. and Rondinelli, C. (2014). Retail market structure and consumer prices in the euro area. European Central Bank Working Paper No. 1744.

Clark, T. S. and Linzer, D. A. (2015). Should I use fixed or random effects? *Political Science Research and Methods* 3(2): 399–408.

Consumers International. (2012). The relationship between supermarkets and suppliers: what are the implications for consumers? http://www.europe-economics.com/publications/the_relationship_between_supermarkets_and_suppliers.pdf. Accessed 25 March 2018.

Correia, S. 2017. Linear Models with High-Dimensional Fixed Effects: An Efficient and Feasible Estimator. Working Paper. http://scorreia.com/research/hdfe.pdf

Cotterill, R. W. (1986). Market power in the retail food industry: evidence from Vermont. *The Review of Economics and Statistics* 68(3): 379–386.

Cotterill, R. W. (1999). Market power and the Demsetz quality critique: an evaluation for food retailing. *Agribusiness* 15(1): 101–118.

Dafny, L., Duggan, M. and Ramanarayanan, S. (2012). Paying a premium on your premium? Consolidation in the US health insurance industry. *The American Economic Review* 102(2): 1161–1185.

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

Dobson, P. W., Waterson, M. and Davies, S. W. (2003). The patterns and implications of increasing concentration in European food retailing. *Journal of Agricultural Economics* 54(1): 111–125.

Eppi, M.J. and Shilling, J.D (1996). Changing economic perspectives on the theory of retail location. In: J Benjamin (ed.), *Megatrends in Retail Real Estate*. Boston: Kluwer Academic Publishers, 65–80.

European Parliament. (2016). Competition in the Food Retail Sector. http://www.europarl.europa.eu/RegData/etudes/STUD/2016/578981/IPOL_STU(2016)578981_EN.pdf. Accessed 1 April 2018.

Evans, W. N., Froeb, L. M. and Werden, G. J. (1993). Endogeneity in the concentration–price relationship: causes, consequences, and cures. *The Journal of Industrial Economics* 41(4): 431–438.

Froeb, L. M. and Werden, G. J. (1991). Endogeneity in the Concentration-Price Relationship: Causes and Consequences. US Department of Justice, Antitrust Division, Economic Analysis Group Discussion Paper, 1 July.

Green, G. and Park, J. (1998). New insights into supermarket promotions via scanner data analysis: the case of milk. *Journal of Food Distribution Research* 29(3): 44–53.

Gyourko, J. and Voith, R. (1993). Leasing as a lottery: implications for rational building surges and increasing vacancies. *Real Estate Economics* 21(1): 83–106.

Harris, M., Kaufman, P., Martinez, S. and Price, C. (2002). The U.S. Food Marketing System, 2002. Agricultural Economic Report AER-811, USDA Economic Research Service (ERS). Washington DC: ERS.

Hernandez, M. A. and Torero, M. (2013). Market concentration and pricing behavior in the fertilizer industry: a global approach. *Agricultural Economics* 44(6): 723–734.

Hovhannisyan, V. (2017). A structural model of cost pass-through: the case of the US yogurt retailing. *Empirical Economics* 1–26. https://doi.org/10.1007/s00181-017-1283-9

Hovhannisyan, V. and Bozic, M. (2013). A benefit-function approach to studying market power: an application to the US yogurt market. *Journal of Agricultural and Resource Economics* 38(2): 159–173.

Hovhannisyan, V. and Bozic, M. (2016). The effects of retail concentration on retail dairy product prices in the United States. *Journal of Dairy Science* 99: 4928–4938.

Hovhannisyan, V., Stiegert, K. W. and Bozic, M. (2014). On the endogeneity of retail markups in an equilibrium analysis: a control-function approach. *Journal of Agricultural and Resource Economics* 39(2): 188–200.

Iposos, M. (2007). Star Power: The Growing Influence of Store Brands in the U.S. Private Label Manufacturers Association.

Kaufman, P. and Handy, C. (1989). Supermarket Prices and Price Differences. U.S. Department of Agriculture, Economic Research Service, Technical Bulletin 1176.

Lamm, R. M. (1982). Unionism and prices in the food retailing industry. *Journal of Labor Research* 3(1): 69–79.

Lin, H. and Wang, I. Y. (2015). Competition and price discrimination: evidence from the parking garage industry. *The Journal of Industrial Economics* 63: 522–542.

Manuszak, M. D. and Moul, C. C. (2008). Prices and endogenous market structure in office supply superstores. *The Journal of Industrial Economics* 56: 94–112.

Marcus & Millichap Real Estate Reports, 2008–2012. http://www.marcusmillichap.com/research/researchreports. Accessed 5 November 2015.

Mourouzi-Sivitanidou, R. (2002). Office rent processes: the case of US metropolitan markets. *Real Estate Economics* 30(2): 317–344.

Muth, M. K., Sweitzer, M., Brown, D., Capogrossi, K., Karns, S., Levin, D. and Zhen, C. (2016). Understanding IRI Household-Based and Store-Based Scanner Data. United States Department of Agriculture, Economic Research Service.

Newman, A. J. and Cullen, P. (2002). *Retailing: Environment and Operations*. London, UK: Cengage Learning EMEA.

Newmark, C. M. (2004). Price-concentration Studies: There You Go Again. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=503522

Okrent, A. and Alston, J. (2012). The Demand for Disaggregated Food-Away-From-Home and Food-At-Home Products in the United States. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2171315

Parente, P. M. and Silva, J. S. (2012). A cautionary note on tests of overidentifying restrictions. *Economics Letters* 115(2): 314–317.

Richards, T. J. and Pofahl, G. (2010). Pricing power by supermarket retailers: a ghost in the machine? *Choices (New York, N.Y.)* 25(2): 1–12.

Rmourouzi-Sivitanidou, R. (2011). Chapter 2: market analysis for real estate. In: D. McKenzie, R. Betts and C. Jensen (eds), *Basic Real Estate Economics*. Mason, OH: Cengage Learning, 30–71.

Rosenblum, C. (2014). Willard Bishop. The Future of Food Retailing. http://willardbishop.com/wp-content/uploads/2014/03/Five-Ways-to-Differentiate-Retail.pdf. Accessed 23 March 2018.

Schmalensee, R. (1989). Chapter 16: inter-industry studies of structure and performance. *Handbook of Industrial Organization*, Vol. II, North-Holland, Amsterdam: Elsevier B. V., 951–1009.

Segal, D. and Srinivasan, P. (1985). The impact of suburban growth restrictions on US housing price inflation, 1975–1978. *Urban Geography* 6(1): 14–26.

Singh, V. and Zhu, T. (2008). Pricing and market concentration in oligopoly markets. *Marketing Science* 27(6): 1020–1035.

Sivitanidou, R. and Sivitanides, P. (2000). Does the theory of irreversible investments help explain movements in office–commercial construction? *Real Estate Economics* 28(4): 623–661.

Steiner, R. L. (2004). The nature and benefits of national brand/private label competition. *Review of Industrial Organization* 24(2): 105–127.

Stock, J. H. and Yogo, M. (2005). Testing for Weak Instruments in Linear IV Regression. In: D.W.K Andrews and J. H. Stock (eds), *Identification and Inference for Econometric Models: Essays in Honor of Thomas Rothenberg*. Cambridge and New York: Cambridge University Press, 80–108.

U.S. Census Bureau. (2012). Annual Retail Trade Survey. https://www.census.gov/retail/arts/how_surveys_are_collected.html. Accessed 7 November 2015.

U.S. Department of Agriculture, Economic Research Service (USDA ERS). (2017a). Global Food Industry. https://www.ers.usda.gov/topics/international-markets-trade/global-food-markets/global-food-industry/. Accessed 3 April 2018.

U.S. Department of Agriculture, Economic Research Service (USDA ERS). (2017b). Retail Trends. https://www.ers.usda.gov/topics/food-markets-prices/retailing-wholesaling/retail-trends/. Accessed 4 April 2018.

U.S. Department of Commerce, Bureau of Economic Analysis. (2015). Population and Per Capita Income by Metropolitan Areas. http://bea.gov/iTable/iTable.cfm?reqid=70&step=1&isuri=1&acrdn=5#reqid=70&step=1&isuri=1. Accessed 23 February 2016.

U.S. Department of Justice. (2011). Agriculture and Antitrust Enforcement Issues in Our 21st Century. https://www.justice.gov/atr/events/public-workshops-agriculture-and-antitrust-enforcement-issues-our-21st-century-economy-10. Accessed on 14 September 2015.

Ward, M. B., Shimshack, J. P., Perloff, J. M. and Harris, J. M. (2002). Effects of the private-label invasion in food industries. *American Journal of Agricultural Economics* 84(4): 961–973.

Wood, S. (2013). Revisiting the US food retail consolidation wave: regulation, market power and spatial outcomes. *Journal of Economic Geography* 2(1): 299–326.

Zhu, T., Singh, V. and Manuszak, M. D. (2009). Market structure and competition in the retail discount industry. *Journal of Marketing Research* 46(4): 453–466.

Table A1. Summary statistics for variables used in empirical analysis

Variable	Mean	SD	Min	Max
Price (US $)	0.051	0.121	3.42e-06	14.333
Convenience (US $Million)	1,891	1,153	186	3,980
Dollar (US $Million)	144.1	75.1	6.3	283
Drug (US $Million)	1,318	846	121	254
Grocery (US $Million)	2,720	1,244	242	5,300
Mass merchandisers (US $Million)	620	284	92	1,111
HHI	683	218	469	1,328
Population (1,000 people)	5,385	5,012	1,217	19,800
Income ($1,000)	43.2	5.6	34.6	58.9

Sources: IRI Retail Data, US Department of Commerce Market Characteristics Data, 2008–2012.
Note: Based on the full sample of 7,755,166 observations.

Appendix: Detailed data description

IRI retail food sales data: IRI collects information on all items scanned at cash registers from more than 11,000 US grocery stores on a weekly basis. The data are then scaled up to reflect sales from stores with annual revenues of $2 million and higher. The IRI dataset contains information on dollar sales and physical volumes for a large group of food products from five departments (dairy, deli, bakery, frozen food and fresh produce) at brand, UPC or item level (a total of 36 billion observations). Most stores in this dataset belong to a retail chain. The remaining non-chain/independent stores are chosen by IRI based on the random stratified sampling method. Specifically, a fraction of stores is dropped each month and replaced by others using a rotating panel design (see Ward *et al.* (2002) for further details).

Given the detailed nature of the data provided, only 41 per cent of the stores listed in the Economic Census were found in IRI in 2012 (Muth *et al.*, 2016). The TDLinx data, on the other hand, provide a much broader coverage of retail stores. For that reason, to calculate our concentration estimates, we utilise TDLinx data as described in the following subsection.

Nielsen TDLinx store characteristics data: We utilise Nielsen TDLinx data on store characteristics to calculate revenue-based HHI estimates of concentration for the respective retail markets. Despite retail grocery competition being limited to local geographic areas, market delineation remains a challenging task (Biscourp, Boutin and Vergé, 2013). In practice, markets are defined based on the competing stores located within a certain radius (Barros, Brito and de Lucena, 2006). In this study, we assume metropolitan areas and/or cities represent retail markets located in different geographic areas of the USA. Given the large number of food products (129) and retail stores (3,144) in our sample, we confine our empirical analysis to 16 US metropolitan statistical areas (Table 1). For identification purposes, we select retail markets with the highest annual variability of the number of retailers (i.e. retailer

Table A2. Descriptive statistics for the market-level HHI estimates, 2008–2012

Market	State	Mean	SD	Min	Max	CV (%)
Charlotte	NC	624	36	574	670	5.77
Chicago	IL	532	19	509	562	3.57
Cincinnati	OH	1,053	30	1,003	1,078	2.85
Columbus	OH	631	20	606	655	3.17
Dallas	TX	525	38	486	574	7.24
Houston	TX	518	32	469	559	6.18
Indianapolis	IN	698	27	661	736	3.87
Jacksonville	FL	759	56	665	801	7.38
Louisville	KY	1,194	74	1,079	1,267	6.20
Milwaukee	WI	721	32	666	744	4.44
Minneapolis	MN	579	13	557	590	2.25
New York	NY	567	31	523	604	5.47
Phoenix	AZ	685	21	649	701	3.07
Sacramento	CA	503	19	478	522	3.78
San Antonio	TX	1,242	58	1,170	1,328	4.67
San Diego	CA	628	65	567	705	10.35

Source: Author calculations based on Nielsen TDLinx Data, 2008–2012.

entry and exit) in the sample period. The retail store formats considered include convenience stores, dollars stores, drug stores, grocery stores and mass merchandisers. Convenience stores tend to be located near consumers with considerably higher prices relative to supermarkets, and normally have a selling area of about 3,000 square feet. The typical convenience store offerings include groceries, soft drinks, snacks and other everyday items. Dollar stores are also known as variety stores that carry a wide range of generic brands or private labels of inexpensive food and drinks, as well as garden tools, personal hygiene and other household consumables. Dollar store prices tend to fall between those at traditional supermarkets and convenience stores. Drug stores are primarily retail stores that offer groceries, cosmetics, books and magazines in addition to pharmaceutical products. Many of these stores rely on non-pharmaceutical products for their revenues. Grocery stores are large self-service supermarkets with a selling area of 4,000 to 27,000 square feet; carry a variety of food and household products; are located in city centres or outskirts; provide convenient shopping hours and have relatively more affordable prices. Finally, mass merchandisers normally have a selling area of up to 100,000 square feet, offer staple goods sold in high volume and quick turnover for less than conventional prices.

As can be observed from Table A1, markets vary widely in terms of market concentration; however, most markets appear to be low to moderately concentrated. The underlying reason may be the way the markets are defined. Further, spatial variation seems to outweigh the temporal variation with the coefficient of variation ranging from as low as 2.25 per cent for Minneapolis, MN to 10.35 per cent for San Diego, CA. To illustrate the contribution of

Table A3. Parameter estimates and standard errors from OLS and IV-OLS regressions

	OLS					IV-OLS				
	2008	2009	2010	2011	2012	2008	2009	2010	2011	2012
Convenience	**2.3661**	**2.5474**	**2.7139**	**2.8626**	**2.9294**	**2.3562**	**2.5434**	**2.7126**	**2.8920**	**2.9056**
	(0.0092)	*(0.0092)*	*(0.0092)*	*(0.0092)*	*(0.0092)*	*(0.0094)*	*(0.0093)*	*(0.0093)*	*(0.0093)*	*(0.0093)*
Dollar	**0.9509**	**1.0377**	**1.1464**	**1.2496**	**1.2155**	**0.9485**	**1.0383**	**1.1480**	**1.2393**	**1.2240**
	(0.0084)	*(0.0083)*	*(0.0083)*	*(0.0083)*	*(0.0083)*	*(0.0084)*	*(0.0084)*	*(0.0084)*	*(0.0084)*	*(0.0084)*
Drug	**2.0705**	**2.2023**	**2.3007**	**2.4995**	**2.5981**	**2.0687**	**2.1946**	**2.2962**	**2.5272**	**2.5801**
	(0.0081)	*(0.0081)*	*(0.0081)*	*(0.0081)*	*(0.0081)*	*(0.0082)*	*(0.0082)*	*(0.0081)*	*(0.0081)*	*(0.0081)*
Grocery	**−0.0456**	0.0019	**0.1048**	**0.2924**	**0.3258**	**−0.0367**	0.0051	**0.1077**	**0.2294**	**0.3693**
	(0.0088)	*(0.0087)*	*(0.0087)*	*(0.0087)*	*(0.0087)*	*(0.0089)*	*(0.0088)*	*(0.0088)*	*(0.0091)*	*(0.0090)*
Mass merchandisers	Ref.	Ref.	Ref.	Ref.	Ref.	Ref.	Ref.	Ref.	Ref.	Ref.
HHI	0.0048	**−0.0086**	0.0047	**0.0116**	0.0062	**0.2308**	**0.2106**	**0.2138**	**0.0319**	**0.3319**
	(0.0035)	*(0.0036)*	*(0.0036)*	*(0.0036)*	*(0.0037)*	*(0.0146)*	*(0.0143)*	*(0.0144)*	*(0.0102)*	*(0.0175)*
Population	**0.2541**	**0.2639**	**0.2779**	**0.2741**	**0.2759**	**0.7139**	**0.7336**	**0.7446**	**0.6586**	**0.7441**
	(0.0251)	*(0.0252)*	*(0.0251)*	*(0.0251)*	*(0.0249)*	*(0.0420)*	*(0.0426)*	*(0.0426)*	*(0.0407)*	*(0.0418)*
Income	**0.0567**	**0.0335**	**0.0388**	**0.0366**	**0.0170**	**0.2246**	**0.1730**	**0.1722**	**0.1205**	**0.1804**
	(0.0030)	*(0.0029)*	*(0.0027)*	*(0.0027)*	*(0.0026)*	*(0.0098)*	*(0.0075)*	*(0.0072)*	*(0.0059)*	*(0.0080)*
Additional controls	(Year × month, product × year, product × channel)					(Year × month, product × year, product × channel)				
Number of observations	1,457,861	1,525,662	1,542,870	1,603,345	1,625,428	1,457,861	1,525,662	1,542,870	1,603,345	1,625,428
First-stage *F*-statistics[a]	*NA*					122.4				

Notes: Values in bold identify parameter estimates statistically different from 0 at the 0.05 level of significance. Store-level cluster-robust standard errors are italicised. The final sample includes 7,755,166 observations.

[a]Presents the Stock–Wright *F*-statistic of zero coefficient on new retail space construction.

Table A4. Parameter estimates and standard errors from fixed-effects and fixed-effects IV regressions that include product × retail format interactions

	Fixed-effects					IV fixed-effects				
	2008	2009	2010	2011	2012	2008	2009	2010	2011	2012
Convenience	Ref.	**0.1760**	**0.3172**	**0.4573**	**0.5027**	Ref.	**0.2428**	**0.2553**	**0.3061**	−0.1496
		(0.0100)	*(0.0169)*	*(0.0258)*	*(0.0382)*		*(0.0411)*	*(0.0919)*	*(0.1090)*	*(0.2621)*
Dollar	Ref.	**0.0746**	**0.1812**	**0.3059**	**0.3135**	Ref.	0.0199	0.0026	−0.0168	−0.3403
		(0.0090)	*(0.0164)*	*(0.0255)*	*(0.0377)*		*(0.0337)*	*(0.0642)*	*(0.1084)*	*(0.2037)*
Drug	Ref.	**0.0718**	**0.1713**	**0.3813**	**0.4683**	Ref.	0.0266	−0.0192	−0.0110	−0.3504
		(0.0088)	*(0.0163)*	*(0.0252)*	*(0.0373)*		*(0.0338)*	*(0.0686)*	*(0.1228)*	*(0.2451)*
Grocery	Ref.	**0.0332**	**0.1588**	**0.3734**	**0.4541**	Ref.	0.0485	0.0221	0.2235	0.2215
		(0.0089)	*(0.0164)*	*(0.0258)*	*(0.0386)*		*(0.0498)*	*(0.0665)*	*(0.1377)*	*(0.3090)*
Mass merchandisers	Ref.	Ref.	Ref.	Ref.	Ref.	Ref.	Ref.	Ref.	Ref.	Ref.
HHI × convenience	**−0.0649**	**−0.1019**	**−0.0885**	**−0.0795**	**−0.0896**	−0.4305	−0.5726	−0.6072	−0.6790	−0.6929
	(0.0213)	*(0.0233)*	*(0.0238)*	*(0.0242)*	*(0.0249)*	*(0.6006)*	*(0.6440)*	*(0.6323)*	*(0.7438)*	*(0.7919)*
HHI × dollar	**0.0492**	**0.0407**	**0.0529**	**0.0583**	**0.0626**	−0.1781	−0.0472	−0.0440	0.1041	−0.0845
	(0.0175)	*(0.0183)*	*(0.0180)*	*(0.0183)*	*(0.0192)*	*(0.1539)*	*(0.1822)*	*(0.1371)*	*(0.2536)*	*(0.1632)*
HHI × drug	**−0.1172**	**−0.1280**	**−0.1213**	**−0.1229**	**−0.1338**	−0.9942	−0.8717	−0.8172	**−0.5419**	**−1.2557**
	(0.0178)	*(0.0185)*	*(0.0186)*	*(0.0184)*	*(0.0188)*	*(0.5827)*	*(0.4902)*	*(0.4832)*	*(0.2384)*	*(0.7701)*
HHI × grocery	**0.0758**	**0.0875**	**0.1214**	**0.1275**	**0.1521**	0.6593	1.0033	0.6861	0.9759	1.2000
	(0.0255)	*(0.0230)*	*(0.0207)*	*(0.0217)*	*(0.0246)*	*(0.5149)*	*(0.7132)*	*(0.4528)*	*(0.6352)*	*(0.7631)*
HHI × mass	−0.0264	−0.0231	0.0107	0.0610	0.0574	**1.7343**	**1.8076**	**1.7441**	**1.8380**	**2.0141**
	(0.0900)	*(0.0967)*	*(0.0980)*	*(0.0972)*	*(0.1027)*	*(0.5839)*	*(0.5641)*	*(0.5374)*	*(0.5622)*	*(0.6561)*
Population	0.1341	0.1428	0.1655	0.1562	0.1598	0.1907	0.2586	0.2648	0.3344	0.2471
	(0.1015)	*(0.1021)*	*(0.1018)*	*(0.1017)*	*(0.1014)*	*(0.3888)*	*(0.4040)*	*(0.4049)*	*(0.3718)*	*(0.3946)*
Income	**0.0478**	**0.0323**	**0.0329**	**0.0429**	**0.0270**	−0.0612	−0.0226	−0.0328	0.0243	−0.0617
	(0.0086)	*(0.0080)*	*(0.0076)*	*(0.0078)*	*(0.0085)*	*(0.1446)*	*(0.1046)*	*(0.1067)*	*(0.0653)*	*(0.1351)*
Additional controls	(Year × month, product × year, product × retail format, store fixed effects)					(Year × month, product × year, product × retail format, store fixed effects)				
Number of observations	1,457,861	1,525,662	1,542,870	1,603,345	1,625,428	1,457,861	1,525,662	1,542,870	1,603,345	1,625,428

Notes: Values in bold identify parameter estimates statistically different from 0 at the 0.05 level of significance. Store-level cluster-robust standard errors are italicised. The final sample includes 7,755,166 observations.

different store formats to the calculation of the HHI estimates, Table A2 presents the distribution of revenue by store format. As can be observed from this table, grocery stores were responsible for an annual average turnover of over $2.7 billion, which accounted for 40.6 per cent of the total revenue generated by all the formats. This was followed by convenience stores ($1.9 billion or 28.2 per cent), drug stores ($1.3 billion or 19.7 per cent), mass merchandisers ($0.6 billion or 9.3 per cent) and dollar stores ($0.1 billion or 2.2 per cent).

US Department of Commerce market characteristics data: Two important descriptors are used to characterise the retail markets under scrutiny. Specifically, we compile population and per capita income data for the period 2008–2012 from the US Department of Commerce, Bureau of Economic Analysis (US Department of Commerce, 2015). Our goal with the inclusion of the population and income variables is to account for the effects of demand-related factors on retail food prices. Markets vary considerably in terms of population density. For example, while the average number of population per square mile is 2,543 in Jacksonville, FL the estimate for New York, NY is 5,684. Further, the coefficient of variation for population change over time varies from as low as 0.49 per cent for Chicago, IL to a high of 2.94 per cent for Houston, TX. Markets also manifest considerable heterogeneity in terms of consumer income. Specifically, average per capita income varies from as low as 36,800 in San Antonio, TX to as high as 56,900 in New York, NY. A general tendency that stands out is that per capita income has declined in a majority of markets following the great recession in 2008. Nevertheless, this effect is predominantly felt in 2009 and starting the following year, income reverted to a rising trend in most markets, eventually surpassing the pre-recession levels.[18]

Marcus & Millichap real estate data: Marcus & Millichap is a commercial real estate brokerage firm founded by George Marcus in Palo Alto, CA in 1971. With a presence in more than 70 cities throughout the USA, it is also one of the largest US companies specialising in real estate investment services. Marcus & Millichap relies on the expertise and local market knowledge of its real estate agents to provide its potential investors with timely and accurate informational basis for efficient and effective decision-making. Moreover, Marcus & Millichap conducts research on the economic and demographic aspects of commercial real estate markets and keeps track of changes to major real estate indicators. The analysis is based on exclusive and comprehensive annual survey data that cover 47 geographic markets throughout the USA and vary by property type.[19] Basic market research results are consequently published in over 1,000 reports annually. *National Retail Report* is one such publication containing annual data on employment, retail sales, asking retail rent, vacancy and retail completions for the

18 Further details are not presented to preserve space, but are available from the authors upon request.

19 The 47 cities included in the survey conform to the IRI Metropolitan Statistical Areas.

geographic markets in question (Marcus & Millichap, 2008–2012). It further ranks the retail markets based on the *National Retail Index (NRI)* computed by the research services division of Marcus & Millichap using cumulative weighted-average scores for various indicators such as retail vacancy rate, new construction, asking rent and employment rate.