Concentration in Product Markets[*]

C. Lanier Benkard[†]
Ali Yurukoglu[‡]
Anthony Lee Zhang[§]

April 2021

Abstract

This paper uses new data to reexamine trends in concentration in U.S. markets from 1994 to 2019. The paper's main contribution is to construct concentration measures that reflect narrowly defined consumption-based product markets, as would be defined in an antitrust setting, while accounting for cross-brand ownership, and to do so over a broad range of consumer goods and services. Our findings differ substantially from well established results using production data. We find that 42.2% of the industries in our sample are "highly concentrated" as defined by the U.S. Horizontal Merger Guidelines, which is much higher than previous results. Also in contrast with the previous literature, we find that product market concentration has been *decreasing* since 1994. This finding holds at the national level and also when product markets are defined locally in 29 state groups. We find increasing concentration once markets are aggregated to a broader sector level. We argue that these two diverging trends are best explained by a simple theoretical model based on Melitz and Ottaviano (2008), in which the costs of a firm supplying adjacent geographic or product markets falls over time, and efficient firms enter each others' home product markets.

Keywords: Concentration, Product markets

JEL classifications: L11, L40, D43

1 Introduction

Industry concentration measures are a key input used in antitrust enforcement, and a barometer that many economists employ for assessing the level of competition in a market. A prominent and growing literature has documented economy-wide increases in industry concentration in the U.S. in the last three decades.[1] Increasing concentration has been linked to declining labor and capital shares,[2] declining investment and productivity growth,[3] and rising markups.[4] This paper uses new data to reexamine these trends. The central innovation in the paper is to construct concentration measures that reflect narrowly defined product markets as would be defined in an antitrust setting, while accounting for cross-brand ownership, and to do so over a broad range of consumer goods and services and a long time frame.

The evidence for broad-based increases in concentration is well established. The most widely cited evidence comes from establishment-level data from the U.S. Economic Census.[5] Similar trends have been demonstrated using firm-level data for public firms from Compustat.[6] The perception of broad-based increases in concentration is also commonplace among politicians and in the popular press.[7]

However, as outlined in detail in Shapiro (2018), there are many problems with drawing antitrust conclusions from the Census data. For antitrust purposes, economists are concerned with the ability of firms to raise prices. Antitrust markets are thus defined based on product substitutability for consumers, using own and cross price elasticities. In contrast, the Census lumps products together that are physically similar and that are produced using similar processes, anywhere in the U.S. A good example of the

---

[1]Peltzman (2014), CEA (2016), Barkai (2016), Gutiérrez and Philippon (2017), Grullon, Larkin and Michaely (2019), Autor et al. (2020), Ganapati (2020), Covarrubias, Gutiérrez and Philippon (2020)

[2]Autor et al. (2020), Barkai (2016)

[3]Gutiérrez and Philippon (2017)

[4]CEA (2016), Barkai (2016), Grullon, Larkin and Michaely (2019), De Loecker, Eeckhout and Unger (2020)

[5]Peltzman (2014), CEA (2016), Economist (2016), Barkai (2016), Gutiérrez and Philippon (2017) Grullon, Larkin and Michaely (2019), Autor et al. (2020), Ganapati (2020), Covarrubias, Gutiérrez and Philippon (2020)

[6]Autor et al. (2020), Gutiérrez and Philippon (2017)

[7]See Shapiro (2018) for an excellent discussion. Early examples include CEA (2016) and Economist (2016).

difference in the two definitions is metal cans, glass bottles, and plastic bottles. Since Census industries are defined based on production and not consumption, all metal cans are in the same Census industry, including soda cans, aerosol cans, paint cans, and many others. Meanwhile, all glass bottles are a separate industry, and plastic bottles a third. These groupings do not make sense for antitrust purposes because paint cans are not a substitute for soda cans, but plastic and glass soda bottles are. Census industries also tend to be too broad. Even at the six-digit level, for example, NAICS 325620 contains at least 42 different industries, including after-shave, deodorant, mouthwash, cosmetics, sunscreen, and hair dye. NAICS 336120 includes all of heavy trucks, buses, garbage trucks, tractors, fire engines, and motor homes.[8] Finally, as noted in Rossi-Hansberg, Sarte and Trachter (2020), Census industries are defined nationally, but many products are delivered locally and are not transportable. Cable TV is a good example in which national concentration has increased dramatically over the past few decades, but this is misleading because local concentration, the relevant statistic for assessing market power to consumers, has decreased just as dramatically. All of these issues are even more present in the Compustat data, which only covers public firms.

Peltzman (2014) says, "One clear question for further research is whether concentration in economic markets has increased... along with the increased concentration in Census Bureau industries." This paper examines exactly this issue. We utilize respondent level data for 1994-2019 from an annual consumer survey available from MRI-Simmons (MRI). The MRI data reports consumers' brand choices across 466 product markets, representing both goods and services.

We have already described some of the difficulties with measuring market concentration across many product markets and a long time horizon in a consistent and meaningful way. The U.S. Horizontal Merger Guidelines suggest identifying the smallest market within which a hypothetical monopolist could impose a "small but significant non-transitory increase of price" (SSNIP). Such an exercise requires a detailed analysis of product level data on quantities and prices over time, and would be extremely costly to implement across such a large number of markets. Instead this paper employs markets

---

[8]Chad Syverson pointed out to us that the trailing zero in NAICS codes 325620 and 336120 indicates that these codes are equivalent to five-digit codes. These are the narrowest industry definitions available in NAICS for these industries, a common issue in NAICS.

defined by a prominent market research data firm whose data are widely used in industry. The market definitions seem close to what might result in an antitrust setting (more details below). Because the survey data contain location data for each consumer, we are also able to measure concentration in geographic sub-markets, an important distinction for products that are delivered and consumed locally.

Another difficulty in measuring concentration across many markets and such a long time period is accounting for joint corporate parent ownership of brands. Many firms own multiple brands in a given product market, and brand ownership changes over time with corporate divestiture and M&A activity, so measuring corporate brand ownership is important to accurately estimate the levels and time trends of product market concentration. We solve this problem by merging the MRI survey data with newly assembled data on brand ownership over time.

These data lead to several interesting findings. Figure 1 presents the median HHI concentration measure over time for four market definitions that differ in their level of geographic and product aggregation. Our central finding is that we document a *decrease* in median product market concentration across a broad range of goods and services since 1994. This result lies in distinct contrast to the findings from production data in the Census. The level of concentration is higher when accounting for the geographic location of consumers in 29 state-groups, but decreases at a similar rate to national concentration. The latter result confirms the main finding in Rossi-Hansberg, Sarte and Trachter (2020), who show that local market concentration has decreased in the establishment data from NETS. However, Rossi-Hansberg, Sarte and Trachter (2020) find increased national concentration in the NETS data even for the most narrow industry classifications, in contrast to our findings.[9]

When product markets are aggregated into broader sectors, our findings reverse: we find increases in concentration over time. After accounting for geographic location at the state level, the rise in concentration at the sector level is small. We find little evidence of firms entering adjacent geographies, as can be seen by the fact that the trends in local and national HHIs are nearly identical. Instead, the joint finding of rising sector concentration and decreasing product market concentration implies that firms are expanding into

---

[9]Smith and Ocampo (2021) also document increasing concentration at both the local and national levels in retail markets.

Figure 1: Median HHI over time, by market definition
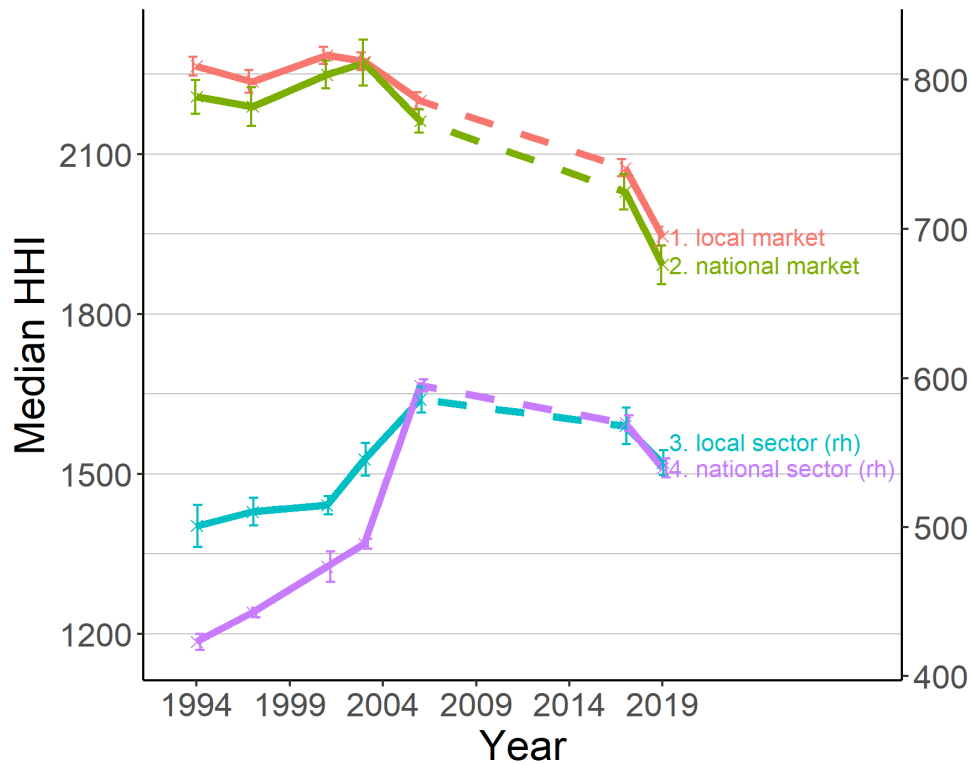
Notes. Local markets are defined as product markets in each of 29 state groups. Sectors are defined by aggregating related national product markets. Product market measures are on the left hand side axis. Sector level measures are on the right hand side axis. Error bars are 95% CIs, based on standard errors from a nonparametric bootstrap.

adjacent product markets within the same sector. Our sector level measurements are more consistent with the results in existing work using establishment-level data,[10] likely because the sector level of market aggregation matches the establishment data more closely.

While we find broad-based decreases in concentration over time, using our market definitions, concentration levels are much higher than in the establishment-level data. In our data the median HHI over all periods is 2180, with 42.2% of industries having an HHI above 2500, the level that is considered "highly concentrated" in the U.S. Horizontal Merger Guidelines. Thus, using this (admittedly too simple) antitrust screen, our data suggests that market power is potentially much higher than previously thought in a large fraction of U.S. product markets. Applying the same simple antitrust screen to the Census data would lead to the opposite conclusion that market power is generally low in U.S. product markets. Autor et al. (2020) report average HHIs from the Census ranging from a low of about 85 in the Services sector in 1987 to a high of 950 in manufacturing in 2007. Even the stricter pre-2010 merger guidelines labeled all of these as "unconcentrated".[11] Our findings on concentration levels parallel Affeldt et al. (2021), who show that concentration levels are much higher than in production data using a sample of market-years in Europe that experienced a merger investigation by the European Commission.

However, while we find high concentration levels, we reiterate that we find no evidence that market power has been getting worse over time in any broad-based way. On the contrary, concentration in the most concentrated industries has fallen as fast as the median industry. In our data the number of industries in the "highly concentrated" range fell from 44.4% in 1994 to 36.6% in 2019. This finding is particularly interesting because it contradicts the prevailing popular opinion (Shapiro, 2018). We speculate that popular perception may be driven by a few prominent large firms, such as Facebook, Apple, Amazon, and Google, which have grown enormously in recent years, as well as high profile mergers in industries such as hospitals (Gaynor, 2018) and airlines.[12]

---

[10]Grullon, Larkin and Michaely (2019), Barkai (2016), Autor et al. (2020), Covarrubias, Gutiérrez and Philippon (2020)

[11]In 2010 the "unconcentrated" range was raised from <1000 to <1500 to reflect practice as detailed in Shapiro (2010).

[12]The US domestic air travel industry saw mergers between US Airways and America West in 2005, Delta

We employ a simple theoretical model based on Melitz and Ottaviano (2008) to explain our main findings. We show that the only force in the model capable of explaining both the product market and sector trends is a reduction in "trade costs", the costs of a firm supplying adjacent product markets. Declining trade costs is also consistent with the well documented increases in corporate profits reported in BEA data[13] and Barkai (2016). The reduction in trade costs could be driven by increasing similarity of production processes, or an increasing importance of logistics and distribution in the production process. As trade costs decrease, efficient firms enter and compete in each others' "home" product markets. For example, in our data Unilever, Proctor and Gamble, and Johnson and Johnson now compete in a vast array of consumer health and cleaning products.

An important implication of our model is that these effects are welfare improving. While sector level concentration increases, the increase is driven by efficiency considerations and consumers benefit. Autor et al. (2020) describe an alternative model in which technological change directly favors more efficient firms. Changes in industry concentration are similarly driven by increasing efficiency and also yield improving welfare. However, their model implies increasing concentration at both the product market and sector levels, whereas our model predicts decreasing concentration at the product market level.

Given the novelty of our data and the contrast between our results and the Census data, we have attempted to verify the external and internal validity of our findings. A weakness of the MRI data is that it is focused on consumer facing product markets, including some services. Purely intermediate goods are largely missing. To evaluate the extent to which our findings are driven by market coverage, we compare our results to those from a subsample of the Census data that is industry matched to the MRI data. Concentration in the subsample has the same overall trend as that in the complete Census data, suggesting that our different findings are due to market definitions and not market coverage. Another weakness of the MRI data is that they are based on surveys rather than actual transactions. To validate the survey data we compared the MRI data with detailed data from industry sources for two industries: airlines and automobiles. For

_____

and Northwest in 2008, United and Continental in 2010, Southwest and Airtran in 2011, American and US Airways in 2014, and Alaska Airlines and Virgin America in 2018.

[13]https://fred.stlouisfed.org/series/W273RE1A156NBEA

both industries concentration in the MRI data closely matches the industry sources in both levels and trends, which gives us confidence that our results are not driven by idiosyncrasies in the survey.

Another potential issue is the product market definitions. Product markets in the MRI data are defined to suit the needs of the client firms who purchase data on their competitors' sales. Close inspection of the data yields the conclusion that the market definitions appear to be reasonable facsimiles of what might result from an antitrust proceeding. The only feature that stands out is that a few of the market definitions are quite narrow. For example, domestic and imported beer are different product markets in the MRI data, as are diet and regular sodas. If the MRI market definitions are too narrow then that could bias us toward finding higher concentration levels overall, but it is not obvious that it would systematically affect the estimates of trends. Moreover, it seems plausible that formal antitrust proceedings might result in these narrow market definitions.

Finally, our model rules out changes in market power due to changes in vertical relationships. It is theoretically possible that large firms like Johnson and Johnson can extract more rents from stores like CVS/Walgreens/Safeway in bargaining if they supply a broader range of products, even if those products are in unrelated product markets (Dafny, Ho and Lee (2019)). If this effect is helping drive the observed trends in market and sector concentration, then there would be an additional force causing market power to increase that is unaccounted for in our model. In that case the welfare effects of the increase in sector concentration would depend on which force is quantitatively more important.

2 Data

2.1 Extracting product information from MRI-Simmons survey

We use respondent level data from the annual "Survey of the American Consumer" available from MRI-Simmons, a market research firm.[14] We use data from 1994 to 2019. MRI surveys approximately 25000 consumers per year in a rolling fashion.

From the survey, we extract all questions that ask consumers to report brands that they purchase. For example, under "Motor oil" in the 2006 survey, the MRI data allows consumers to report purchases of 24 different brands of motor oil, such as Valvoline, Castrol, Amoco, Havoline, and Chevron, as well as an "Other" option. In total, we extract brand purchase information for 337 products; we will call these "product markets". We divide these product markets into 17 broader groups, such as "Home products – Food" or "Airlines"; we will call these broader groups "sectors". Table 1, which we describe below, lists all the sectors in our data, the number of product markets in each sector, and examples of product markets within each sector. We also distinguish between "manufacturing" and "non-manufacturing" sectors. The manufacturing sectors tend to have a larger number of product markets.

In addition to brand purchase information, the survey asks respondents for demographic information, in particular, the state group that respondents live in. There are 29 state groups; large states are reported separately, but some less populated states that are close together are grouped together, such as Minnesota/Iowa, Nebraska/Kansas, Arkansas/Louisiana/Oklahoma. We use state group information so that we can calculate product purchases at the level of stategroup-markets. Further details of data cleaning are described in appendix A.

MRI data are well known in industry and commonly employed in media planning. Gentzkow and Shapiro (2011) use the MRI data to measure ideological segregation in news consumption. Crawford and Yurukoglu (2012) use the MRI data to estimate demand for cable television services. Bertrand and Kamenica (2018) use the MRI to document similarity in consumption between different demographic groups over time.

---

[14]The firm administering the survey has undergone several changes in ownership and has been previously known as Mediamark Research Inc (MRI) and GfK MRI.

2.2 Brand ownership information from Kantar Adspender

We derive brand ownership information by merging MRI brand names to Kantar Adspender. Kantar Adspender is a database that tracks brands' advertising expenditures across different advertising media. We digitized hard copies of Kantar Adspender for the years 1992, 1997, 2001, 2003, 2006, and downloaded data from Kantar Adspender in 2017 and 2020.[15] Kantar Adspender contains data on advertising expenditures; the brand name advertised, and the ultimate parent company of the brand. For the pre-2016 data, only a single parent company name is available. For the 2017 and 2020 data, there are a number of different ownership fields: "ultimate parent", "parent", "subsidiary", and "advertiser". We use the "ultimate parent" field.

For each of the years in which we see Kantar Adspender, we merge the corresponding year of the MRI data to Kantar Adspender. The only exceptions are that we merge the 1992 Adspender to the 1994 MRI and the 2020 Adspender to 2019 MRI. We merge the Adspender by brand name using a two-stage fuzzy string-matching algorithm that we describe in detail in appendix A.2. We are able to match over 80% of brands in most sectors to an ultimate parent, and over 90% of market share for all sectors other than pet products (see table 1). The brand matching allows us to tell when different brands are owned by the same ultimate owner. For example, while the 2006 data reports 24 different brands of motor oil, most of these brands are owned by three companies: Chevron-Texaco, Exxon Mobil, and Royal Dutch Shell.

There is a nontrivial amount of brand co-ownership in our data. The average brand owner in our data set owns 2.88 brands. The brand ownership distribution is highly skewed, with 76.289% of owners owning only one brand, whereas the largest brand owner owns 253 brands. Ownership across product markets is also nontrivially large: the average owner owns brands across 2.11 product markets. 26.43% of brand owners own brands across at least 2 markets. Tables 2 and 3 show the largest brand owners for different years, for manufacturing and non-manufacturing separately. For manufactures, some of the largest owners are Procter & Gamble, Kraft Heinz, Unilever, Johnson &

---

[15]Kantar Adspender has historical information about advertising expenditures, but brand ownership information is backfilled: brands advertised in earlier years are assigned to their most recent ultimate owner. Using historical hard copies of Adspender allows us to circumvent this problem. However, the hard copies are unavailable after 2006.

Johnson, and Clorox. For non-manufactures, largest owners include Visa, State Farm, and Blue Shield.

2.3 Computing market shares

The MRI data contains indicators for whether consumers have purchased a given brand, but typically does not provide quantity or expenditure information. As a workaround, we compute market shares assuming that if a customer purchases multiple products in one market, she purchases the same quantity of each product.

Let B_{mo} represent the set of brands owned by owner o, let I_s represent the set of customers living in state s, and let I represent the set of all consumers. The market share of owner o in state s, market m, time t, is:

$$s_{omst} = \frac{\sum_{b \in B_{mo}} \sum_{i \in I_s} e_{ibmt}}{\sum_o \sum_{b \in B_{mo}} \sum_{i \in I_s} e_{ibmt}} \tag{1}$$

where e_{ibmt} is an indicator variable, for whether customer i reports purchasing brand b in market m at time t, multiplied by the sampling weight on customer i. The national market share of owner o in market m, time t, is:

$$s_{omt} = \frac{\sum_{b \in B_{mo}} \sum_{i \in I} e_{ibmt}}{\sum_o \sum_{b \in B_{mo}} \sum_{i \in I} e_{ibmt}} \tag{2}$$

We can also aggregate to the higher level of sectors, which we will index by k. Let M_k represent the set of markets in sector k. The national market share of owner o in sector k, time t is:

$$s_{okt} = \frac{\sum_{m \in M_k} \sum_{b \in B_{mo}} \sum_{i \in I} e_{ibmt}}{\sum_o \sum_{m \in M_k} \sum_{b \in B_{mo}} \sum_{i \in I} e_{ibmt}} \tag{3}$$

Using each of these market shares, we can then compute concentration metrics at the level of stategroup-markets, markets, stategroup-sectors, and sectors.

The MRI data includes a number of choices such as "Other" or "Store brand," that may correspond to multiple brands; treating these as single brands may lead to overestimating market concentration. We take the opposite approach, which is conservative for estimating

concentration levels: we include "Other" and "Store brand" in the denominator when calculating the shares (1), (2) and (3), but do not include them as owners.[16] Essentially, this is like assuming that "Other" and related options constitute a continuum of infinitely small brands.

2.4 Computing HHIs

The HHI is a convex function of market shares, which introduces an upward bias to HHIs calculated using unbiased finite-sample estimates of market shares. To account for this, all HHIs we report in the paper are adjusted using a nonparametric bootstrap procedure to correct for finite-sample bias, which we describe in appendix A.3. The bias adjustment reduces the estimates of stategroup-product market HHIs (the lowest level of aggregation) by around 150 points (out of 10,000), but has negligible effects on HHI estimates at other levels of aggregation.

2.5 Summary of Sectors

Table 1 describes the sectors we analyze. Our main results focus on a balanced panel of the set of product markets that appear in each year from 1994 to 2019. A non-trivial number of product markets appear in only a subset of years. For example, wireless handsets were not measured prior to 2004. In appendix B we also report results separately for the unbalanced panel consisting of all product markets that appear in the data.

We categorize sectors into manufacturing and non-manufacturing. The non-manufacturing sectors are airlines, car rental, financial, hotels, insurance, retail, and restaurants. The data tend to cover many product markets within manufacturing sectors, and fewer within non-manufacturing sectors. Our main results include both the manufacturing and non-manufacturing sectors. For robustness, we also report results separately for the manufacturing and non-manufacturing sectors.

---

[16]In addition, treating "Other" or "Store brand" as single brands increases measured concentration levels slightly but does not have a large affect on measured trends.

Table 1: Descriptive Statistics

Sector	# Markets	# Brands	# Owners	Matched brands %	Matched marketshare %	Example markets
Airlines	2	28.3	20.1	95.7	99.3	DomesticTravelAirlinesused, ForeignTravelAirlinesUsed
Apparel	2	40.9	31.1	89.1	96.3	AthleticShoesBrandsBought, WomensLingerieUndergarments
AutoProducts	15	196.9	97.6	86.7	96.7	AirFilters, WindshieldWipers
Automobile	2	52	29.1	89.5	98.8	AutomobilesAndOtherVehiclesManufacturer, MotorcyclesMake
Beverages	39	531.3	176.9	83.6	95.1	BottledWaterSeltzer, Vodka
CarRental	3	20.6	7	100	100	CarRentalBusinessUse, TruckTrailerRentalCompanies
Electronics	5	78.1	46.3	89.0	96.4	Batteries, TelevisionSetsBrands
Financial	3	42.6	26.9	92.7	99.1	CreditCards, RealEstateWhichAgentUsed
Health	62	1,127.9	239.7	86.8	94.1	AdhesiveBandages, WartRemovers
HoProdChild	17	103.3	31.4	93.6	98.5	BabyBathWashAndSoap, VitaminsForChildren
HoProdFood	126	1,700.7	445	83.0	93.8	AmericanPasteurizedProcessedCheese, Yogurt
HoProdNonfood	43	511.1	170.6	81.3	93.0	AirFreshenersCarpetRoomDeodorizers, WritingInstrumentsBrands
HoProdPets	7	118.3	41.4	76.5	86.6	CannedWetCatFood, PackagedDryDogFood
Hotels	1	34.4	17.1	96.3	98.0	HotelsMotelsWhereStayed
Insurance	4	102.3	56.7	91.4	98.4	AutoInsurance, MedicalInsuranceCompanies
Restaurants	2	114.4	92.7	88.5	94.8	FamilyRestaurantsSteakHouses, FastFoodDriveInRestaurants
Retail	4	74.4	64.7	89.7	97.0	ApplianceHardwareElectronicsStoresTimesShopped, FurnitureStoresTimesShopped

Notes. Summary statistics by sector. All numbers are averaged by year. "# markets" is the number of product markets in the sector. "# brands" and "# owners" are respectively the total number of brands and owners within a sector. "Matched brands %" and "Matched marketshare %" are, respectively, the number of brands and fraction of market share matched to owners. "Example markets" shows examples of markets within the sector

Table 2: Top 10 brand owners by year, manufacturing sectors

rank	1994	2003	2017
1	procter & gamble co	altria group inc	procter & gamble co
2	philip morris cos inc	procter & gamble co	kraft heinz co
3	unilever nv	unilever	unilever
4	conagra inc	general mills inc	general mills inc
5	johnson & johnson	conagra foods inc	johnson & johnson
6	nestle sa	clorox co	conagra brands inc
7	campbell soup co	pepsico inc	clorox co
8	johnson sc & sons inc	johnson & johnson	nestle sa
9	general mills inc	reckitt benckiser plc	sc johnson & son inc
10	clorox co	nestle sa	jm smucker co

Notes. Top 10 largest brand owners by year for manufacturing sectors.

Table 3: Top 10 brand owners by year, non-manufacturing sectors

rank	1994	2003	2017
1	sears roebuck & co	state farm mutual auto	visa usa inc
2	state farm mutual auto	wal-mart stores inc	state farm mutual auto
3	k mart corp	visa usa inc	blue cross & blue shie
4	visa international	home depot inc	home depot inc
5	wal-mart stores inc	blue cross & blue shie	mastercard intl inc
6	pepsico inc	allstate corp	lowes cos inc
7	blue cross & blue shie	mcdonalds corp	wal-mart stores inc
8	southland corp	cendant corp	berkshire hathaway inc
9	mcdonalds corp	mastercard intl inc	allstate corp
10	mastercard internation	ito-yokado co ltd	seven & i holdings co

Notes. Top 10 largest brand owners by year for non-manufacturing sectors.

3 Results

Figure 2 shows the distribution of HHI's in our data at the stategroup-market ("local market"), market, stategroup-sector ("local sector"), and sector level over time. The DoJ-FTC 2010 Horizontal Merger Guidelines define industries with HHI's between 1500 and 2500 as "moderately concentrated," and above 2500 as "highly concentrated." According to the guidelines, mergers that raise the HHI in moderately or highly concentrated industries often warrant scrutiny.[17]

We find much higher concentration levels than those measured using production data. The median HHI in local product markets during the whole period is 2180, with an average of 42.2% of industries falling in the "highly concentrated" range. For comparison, Keil (2017) reports a median HHI of 450 between 1990 and 2012 using data from the Economic Census. Autor et al. (2020) report average HHIs from the Census ranging from a low of about 85 in the Services sector in 1987 to a high of about 950 in manufacturing in 2007. Accounting for brand co-ownership also makes a large difference. While not the main focus of their paper, Neiman and Vavra (2018) report average HHIs of about 30 for categories in the Nielsen scanner data, not accounting for multi-product firms.

Still focusing on local product markets, the most noticeable change in the distribution of HHIs is that concentration has fallen over time – the median HHI fell from 2265 in 1994 to 1945 in 2019. Importantly, concentration fell as much in the most concentrated industries as in the median industry. The 90th percentile HHI fell from 5325 in 1994 to 4570 in 2019, while the 75th percentile fell from 3561 to 3171. Figure 3 shows that the fraction of firms in the "highly concentrated" range fell from 44.4% in 1994 to 36.6% in 2019. Thus, while we find high concentration levels, particularly in 1994, according to our data there has been substantial improvement over time. These findings are in conflict with the prevailing popular opinion that increases in market power in the U.S. have been large and widespread (Shapiro, 2018).

The levels and trends in national product market HHIs are close to local product market HHIs at all percentiles. We infer that there is little evidence in our data of firms entering more local geographic markets over time. Our finding of decreasing local market

---

[17]Nocke and Whinston (2020) demonstrate that changes, rather than levels, in HHI are more informative for unilateral merger effects in commonly used demand and conduct models.

Figure 2: HHI percentiles at different market levels over time
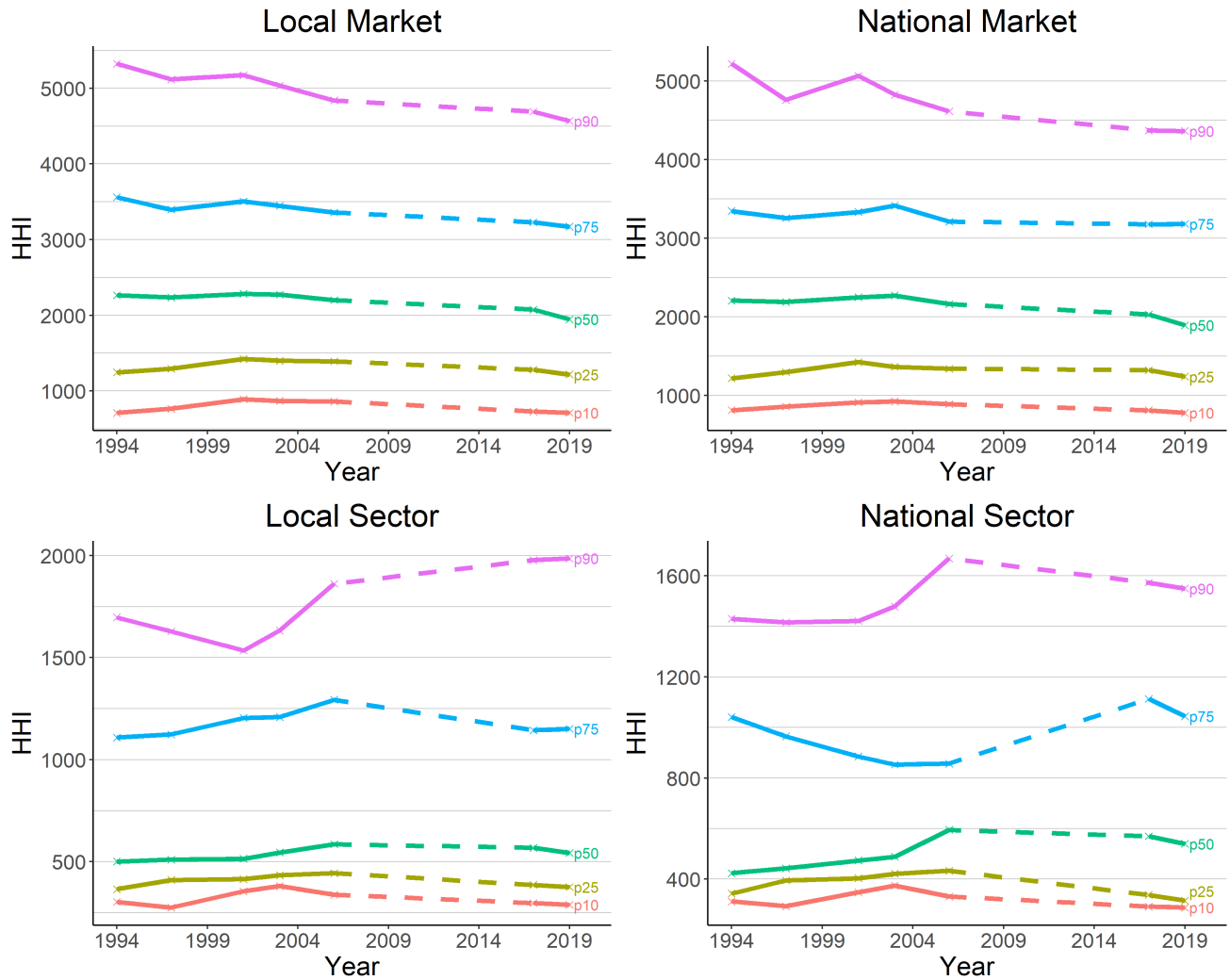
Notes. Percentiles of HHI over time, at the stategroup-product market (top left), product market (top right), state group-sector (bottom left), and sector (bottom right) levels.

Figure 3: Fraction of local markets by concentration
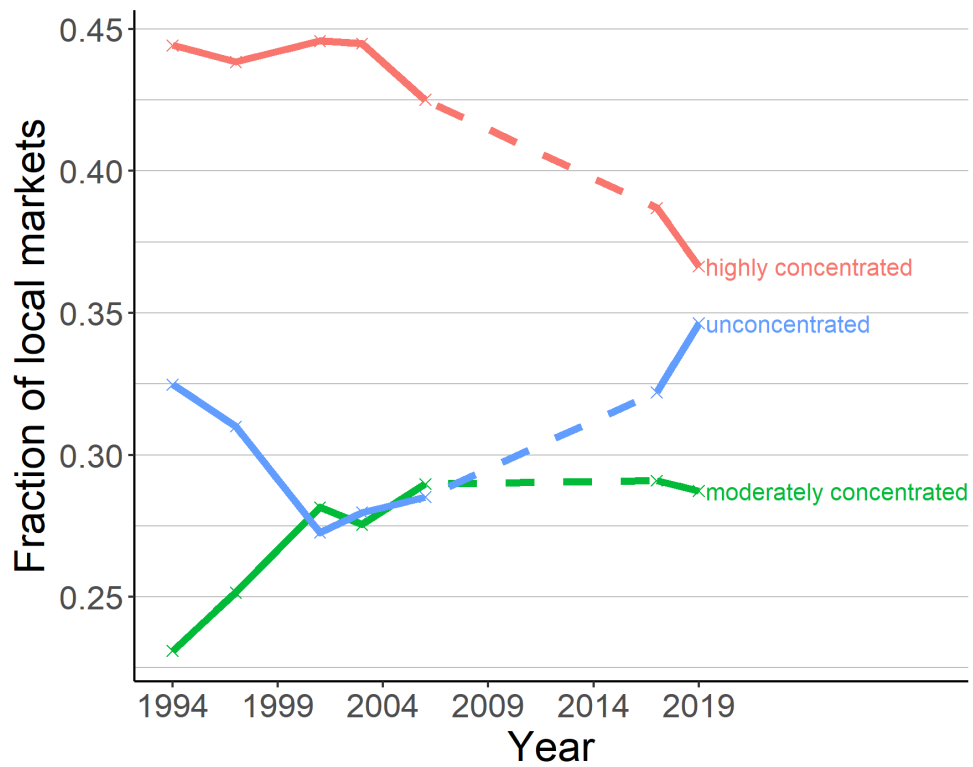
Notes. The fraction of local markets by their level of concentration: highly concentrated (HHI higher than 2500), moderately concentrated (HHI between 1500 and 2500), and unconcentrated (HHI lower than 1500).

concentration confirms the main finding in Rossi-Hansberg, Sarte and Trachter (2020). However, Rossi-Hansberg, Sarte and Trachter (2020) find increasing national concentration, even for the most narrow industry definitions. The differences could be explained by the fact that the NETS data in Rossi-Hansberg, Sarte and Trachter (2020) is measured at the establishment or plant level (point of production),[18] whereas the MRI data is observed at the consumer level (point of consumption). While establishment data would accurately reflect consumer retail competition, for goods that are produced in a small number of plants and sold nationally, establishment data would show a skewed local market share relative to national market share even if local and national consumption market shares were similar.

On the other hand, aggregating product markets into broader sectors presents a qualitatively different story. At this higher level of aggregation, the 50th, 75th, and 90th percentiles all experience clear increases in concentration over time. The difference between product level and sector level HHI is evident at both the local and national levels. The joint finding of rising sector concentration and decreasing product market concentration implies that firms are expanding into "adjacent" product markets within the same sector. While this could be achieved through a combination of mergers and *de novo* entry, in concert with horizontal mergers and exit, our data do not allow us to distinguish the precise mechanisms behind the increase because the data contain numerous name changes for both firms and brands that make it difficult to link them over time. Our sector level measurements are more consistent with the results in existing work on the establishment data,[19] likely because the sector level of market aggregation matches the establishment data more closely.

While our findings above well represent the overall trends for the whole economy, every individual product market and sector is different. Figure 4 shows local market HHI trends by sector. HHI in new automobiles, which account for roughly 3% of consumer expenditure, fell from 2508 to 1322. The largest increase in HHI is in the car rental market, where HHI grew from 1937 to 3677. More generally, we see many manufacturing sectors experiencing a decrease in local market HHI, whereas most non-manufacturing sectors

---

[18]See Crane and Decker (2019) for details on the reliability of the NETS sales data.
[19]Grullon, Larkin and Michaely (2019), Barkai (2016), Autor et al. (2020), Covarrubias, Gutiérrez and Philippon (2020)

experience no substantive change (exceptions being Financial and Car Rental).

Figure 5 plots the local product markets that experienced the largest changes. The markets with the largest HHI decreases generally experienced growth in new brands or a shift of market share to store brands rather than spreading of share among existing brands. For example, in glue, Gorilla Glue entered the market in 1999 and increased its market share to above 30% in 2019, accounting for a large fraction of the decrease in share by the dominant brands Elmer's and Krazy, both owned by the same parent. We also observe the parent of the Gorilla Glue company entering into other product markets such as skin care by 2019. The decrease in concentration in rubber gloves is due to entry by Proctor and Gamble with the introduction of Mr. Clean brand gloves.[20] By 2019, this brand took a significant market share from market leader Playtex brand.

The largest increase markets include increases due to merger activity as well as a concentration of share into the highest selling brands. For example, among the largest increasing markets are car rental, dry cake mixes, and condoms. Doane et al. (2018) documents a series of mergers in the car rental industry. The increase in concentration for dry cake mixes is driven by the 2000 acquisition of Pillsbury by General Mills group.[21] The driver in condoms was through growth of the share of the top brand Trojan during this time period.
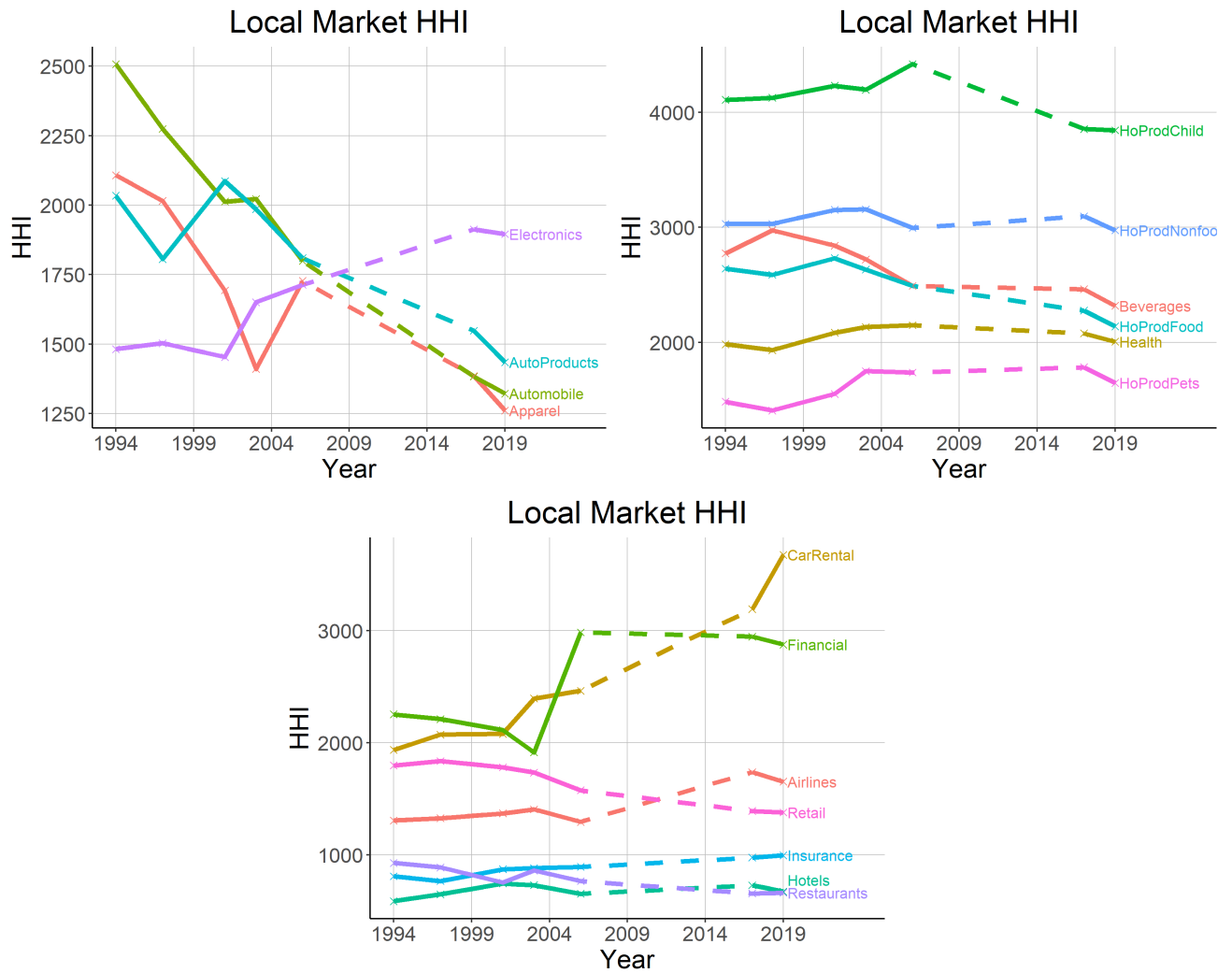
3.1 External validity checks

Given the novelty of our data and the contrast between our results and those from the production data, we have made several attempts to check the external validity of our findings. The biggest threats to external validity are market coverage, market definitions, and the survey nature of the data. We address these issues first.

---

[20]https://www.core77.com/posts/22044/international-home-housewares-show-2012-mr-cleans-new-line-of-cleaning-gloves-22044

[21]https://www.wsj.com/articles/SB963782500794995149

Figure 4: Local market HHI over time, by sector

Notes. HHI over time, at the stategroup-product market level. Each line shows the expenditure-weighted average of HHIs, for all stategroup-markets in a given sector. The left panel shows results for manufacturing, the right panel for food, beverage, and health products, and the bottom panel shows results for non-manufacturing. Appendix figure A.6 replicates this figure for different levels of aggregation.

Figure 5: Largest Changes in local HHI
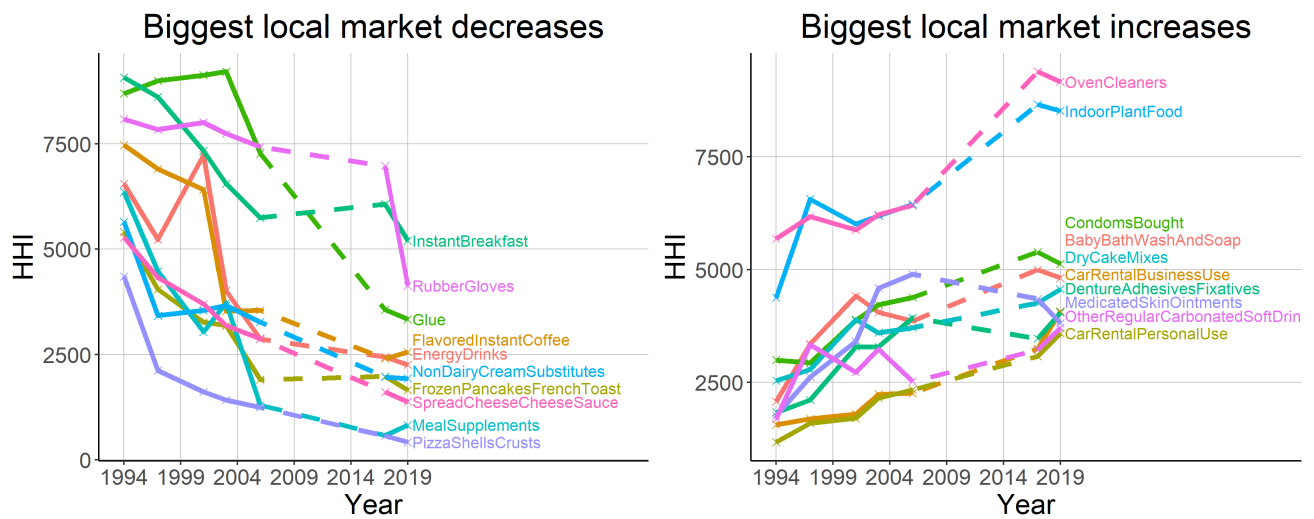
Notes. HHI over time, at state group level, for the 10 product markets with the largest decreases and increases in HHI. Each line shows the expenditure-weighted average of HHIs, for all local markets in a given market.

3.1.1 Market coverage

A number of other papers have shown that in the Census data concentration in production has increased over time even at the six-digit NAICS level. The MRI data covers fewer markets than the Census and the selection is not random because it systematically omits purely intermediate goods, so in this section we check to see if the differences could be driven by sample selection. We recalculate concentration in the Census data, generate a subsample of Census data that best matches the MRI product market sample, and confirm that concentration rises in the subsample of the Census data. We clean the raw Census data following Barkai (2016) to get a year-industry panel for 1997 to 2012 corresponding to 2012 six-digit NAICS codes. Then, we hand-match product markets from MRI to the six-digit NAICS codes. In general, this is a many-to-many mapping. We then use the Census weights to recompute HHIs.

Figure 6 plots changes in C4, which has better availability than HHI in the Census data. We verify the findings of other papers: using the Census six-digit NAICS codes as definitions of product markets, concentration is increasing over time. The matched subsample of markets from MRI has higher C4 than the unmatched markets, but the trend in concentration is the same.[22]

3.1.2 Comparison to industry specific measures

One downside of the survey data is that they do not reflect actual transactions. Respondents may not remember what they purchased or may not face strong incentives to accurately report what they purchased. To check the accuracy of market shares from our dataset, we compare our results to two industry specific datasets.

For automobiles, we use sales information to construct product market HHI using Ward's Automotive Research data as used in Berry, Levinsohn and Pakes (1995). Refer to Grieco, Murry and Yurukoglu (2020) for a detailed description of the cleaning process. The results are shown in the left panel of figure 7. Both the levels and trends are very similar between the two datasets: HHI declines from around 2200-2500 in 1994 to around

---

[22]In appendix B.1 we also use a regression technique that allows us to include in the comparison industries whose NAICS industry definition changed over time, with similar results.

Figure 6: Census concentration
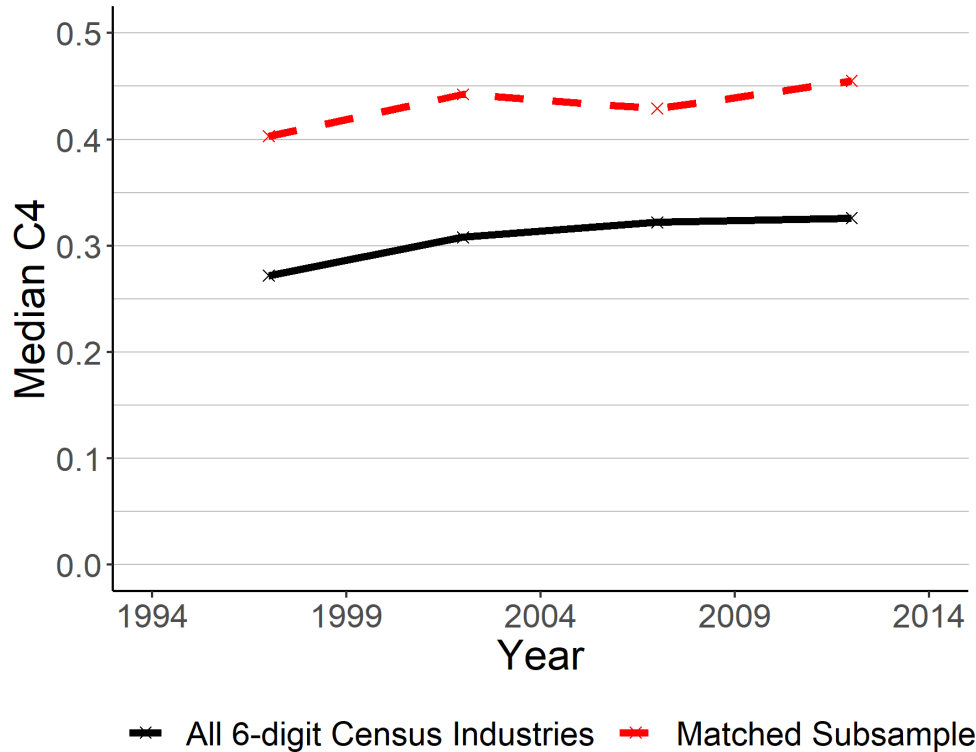
Notes. Median C4s using 2012 Census six-digit NAICS codes. All Census industries (black), and subsample of Census industries matched to MRI (red).

1200 in 2018.

For airlines, we use the Airline Origin and Destination Survey from the Bureau of Transportation Statistics (BTS). This survey is a 10% sample of airline tickets from all US domestic carriers and includes origin, destination and ticket details. We aggregate total revenues by carrier group which include airlines that operate under different brands but under common ownership. Using the BTS data, we can construct both local and national HHI measures. To calculate local HHIs, for comparability to the MRI, we aggregate the BTS data, by total revenue, to the level of MRI state-groups, and then we calculate HHIs at the stategroup level. We define stategroups by origin airport states, but we have verified that the results hold for destination airports states as well. For national HHIs, we aggregate BTS data to the national level to construct market shares by total revenue.

We show the results in the right panel of figure 7. As with the automobile data, the two datasets are very similar in both levels and trends, at both the local and national level.

Figure 7: Automobile and Airlines Robustness Check
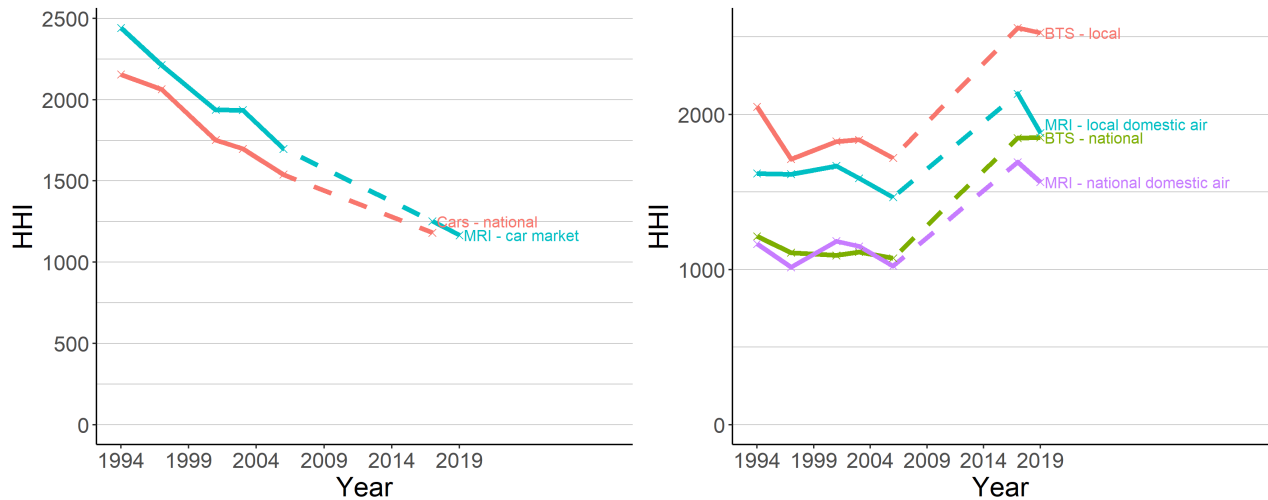
Notes. The left panel shows the estimated automobile market HHI from the MRI (blue), against the estimated national HHI for the car market, from Ward's. The right panel shows the estimated HHI from the MRI, for domestic airlines, at the local market (blue) and national market (purple) levels, against BTS airline data local (red) and national (green) HHIs. For local concentration we take median HHIs.

3.2 Robustness Checks

We have also computed concentration using several alternative concentration measures and market definitions. The details of these analyses can be found in appendix B. Our results hold if we measure concentration using C2 and C4 instead of HHIs. The results reported above are for a balanced panel of industries, but they also hold for the full unbalanced panel and under alternative assumptions about industries that change definition over time. Our results also hold when we re-weight sectors according to expenditure shares from the Consumer Expenditure Survey. We also examine concentration at two intermediate levels of market aggregation between product markets and sectors, and we

find that the general trend continues to hold: concentration is increasing over time at higher levels of aggregation, and decreasing at lower levels of aggregation.

4 Model

To rationalize the empirical results, we use a simple version of the Melitz and Ottaviano (2008) model to study the determinants of concentration at different levels of market definition. We derive analytical expressions for HHIs in the model, and show which changes in model primitives are consistent with the trends we observe in the data.

In the model there are two identical markets, 1 and 2, indexed by j. These markets have two alternative interpretations. They can be thought of as two geographic regions, such as US states. A firm that is headquartered in one state can export to another state, but at a higher marginal cost. Alternatively, the two markets could be different product markets within the same sector, such as orange juice and soda. Each firm specializes in one product (its "home" product market), and can also produce the other product, but at a higher marginal cost.

Each market contains a unit mass of consumers. Preferences are defined over a continuum of differentiated varieties, where each firm, indexed by $i \in \Omega$, produces a single variety. There is a numeraire good, q_0. Consumers' utility is:

$$U = q_0 + \alpha \int_{i \in \Omega} q_i di - \frac{1}{2}\gamma \int_{i \in \Omega} (q_i)^2 di - \frac{1}{2}\eta \left(\int_{i \in \Omega} q_i di \right)^2$$

with a larger γ implying a stronger taste for variety.

We assume that there is an infinite measure of potential firm entrants in each market. Each entrant must pay some irreversible fixed cost f_E to enter. Once a firm i has entered, the firm draws a marginal cost c, distributed as:

$$G(c) = \left(\frac{c}{c_M} \right)^k$$

That is, $\frac{1}{c}$ is Pareto distributed, with lower bound $\frac{1}{c_M}$, and shape parameter $k \geqslant 1$. Firms

24

can produce in their local market at constant marginal cost c. Firms can also "export" to the other market. The marginal cost a firm faces for producing in the export market is τc, where $\tau > 1$. Thus, when τ is low, firms have similar production costs in both markets; when τ is high, a firm based in one market faces a high marginal cost of producing in the other market.

The interpretation of τ differs depending on whether markets represent geographic regions or product markets. If the two markets are interpreted as geographies, τ can be thought of as representing product trade costs, such as physical transportation costs, or increased operational costs of advertising and selling across state borders. If the two markets are interpreted as product markets within a sector, τ can be thought of as the additional costs that a specialized firm faces when it produces a different product. Firms specializing in orange juice might have a higher marginal production cost for soda due to accumulated experience from learning-by-doing, or technological factors like production equipment that is better suited to orange-juice production. The benefits to specialization may also vary across sectors and over time depending on technology and the composition of marginal costs. Certain costs, such as transportation, logistics, and marketing costs, are more likely to be common within a firm across different product markets. In sectors and time periods where these factors make up a larger share of marginal costs, τ would tend to be lower.

In equilibrium, firms enter until the expected profits from entry equal f_E for entrants in both markets. There will be some cost cutoff c_D such that firms who draw costs c higher than c_D will choose to produce nothing. The following proposition characterizes the measure of firms that enter, the cutoff cost for production, and the domestic and export quantities produced by firms in the unique equilibrium of the model.

Proposition 1. *There is a unique equilibrium of the model. The production cost cutoff satisfies:*

$$c_D = \left[\frac{2\gamma \, (k+1) \, (k+2) \, (c_M)^k \, f_E}{\left(1 + \frac{1}{\tau}\right)} \right]^{\frac{1}{k+2}} \tag{4}$$

The number of varieties produced in each market, by domestic producers as well as exporters, is:

$$N = \frac{2(k+1)\gamma}{\eta} \frac{\alpha - c_D}{c_D} \tag{5}$$

In equilibrium, firms produce in their local market if $c < c_D$. The quantity produced by a firm with cost c in the local market j is:

$$q_j(c) = \frac{1}{2\gamma}(c_D - c) \tag{6}$$

Firms export positive quantities if $\tau c < c_D$. The quantity produced by a firm with cost c in the export market j is:

$$q_j^X(c) = \frac{1}{2\gamma}(c_D - \tau c) \tag{7}$$

Consumer welfare is:

$$1 + \frac{1}{2\eta}(\alpha - c_D)\left(\alpha - \frac{k+1}{k+2}c_D\right) \tag{8}$$

We build on Melitz and Ottaviano by calculating market shares and Herfindahl-Hirschman indices (HHIs) at the local market and sector level and showing how they vary with model primitives. Define the total quantity in market j as the integral over all firms' production in market j:

$$Q_j \equiv \int_{i \in \Omega} q_j(c_i)\, di$$

We can then define the market share of a firm i in market j, $s_j(i)$, as the ratio of quantity produced by firm i in market j to the total quantity in market j, and the aggregate market share $s(i)$ as i's total quantity across both markets, divided by aggregate market quantity. That is:

$$s_j(i) \equiv \frac{q_j(c_i)}{Q_j}, \quad s(i) = \frac{\sum_{j=1}^{m} q_j(c_i)}{\sum_{j=1}^{m} Q_j} \tag{9}$$

These are the natural continuous analogs to market shares with a discrete number of firms. We can then define HHIs at the market and aggregate level as follows.

Definition 1. Define the local HHI in market j as:

$$HHI_j = \int_{i \in \Omega} \left(s_j\left(i\right)\right)^2 di \tag{10}$$

and the aggregate HHI as:

$$HHI = \int_{i \in \Omega} \left(s\left(i\right)\right)^2 di \tag{11}$$

As noted above, local and aggregate markets can be interpreted either geographically, as state-group and national market shares, or as product markets and sectors. Since we are primarily interested in the convergence between market and sector HHIs, we will describe local HHIs as "market" and aggregate HHIs as "sector" HHIs. The following proposition characterizes market and sector HHIs in the model.

Proposition 2. *Market HHIs are:*

$$HHI_j = \frac{1}{N} \frac{2 + 2k}{2 + k} \tag{12}$$

Sector HHIs are:

$$HHI = HHI_j -$$
$$\frac{k+1}{N\left(1 + \frac{1}{\tau^k}\right)(k+2)} \left[1 - \frac{1}{2}k\left(k+1\right)\left(k+2\right)\left(\frac{1}{k\tau^k} - \frac{2}{(k+1)\tau^{k+1}} + \frac{2\tau - \tau^2}{(k+2)\tau^{k+2}}\right)\right] \tag{13}$$

The market HHI is strictly increasing in τ. The difference between the market and sector HHIs, $HHI_j - HHI$, is also strictly increasing in τ.

Proposition 2 shows that the market HHI is a function of the number of firms that enter in equilibrium and k, which controls the dispersion in firms' marginal cost draws. Product market HHIs are always increasing in τ. Intuitively, this is because, when τ decreases, productive firms have an easier time competing in markets besides their home market, which always increases product market competition in equilibrium.

From (13), the sector HHI is always lower than the product market HHI. The difference between sector and market HHIs is also increasing in τ. Intuitively, when τ is high many

firms produce in only one market and markets may be quite concentrated, even if the total number of entrants across both markets is fairly large. On the other hand, when τ is low most firms produce similar amounts in both markets, so sector and product market concentration are similar.

4.1 Comparative statics and interpretation of results

Figure 8 shows how changes in model primitives impact market and sector HHIs and consumer welfare. The left column shows the result of varying trade costs, τ. As in proposition 2, when trade costs decrease, market HHIs tend to decrease and sector HHIs tend to increase.[23] When trade costs are low, productive firms enter and compete in more export markets and this additional competition decreases market HHIs. However, entering in more markets also makes firms larger at the sector level, increasing sector HHIs.

The middle and right columns of figure 8 show how HHIs change with other model parameters. The middle column shows that higher entry costs, f_E, tends to increase concentration at both the market and sector levels simultaneously, without a large effect on the difference. The right column shows the effect of varying k, which controls the dispersion in firm productivity draws; lower values of k correspond to more dispersion in productivity.[24] In order to isolate the effect of changing productivity dispersion, as we vary k we simultaneously vary the entry cost, f_E, so that the firm entry cutoff c_D remains fixed. Similar to fixed costs, decreasing k tends to increase both market and sector HHIs simultaneously. Aghion et al. (2019) and Hsieh and Rossi-Hansberg (2019) present related models in which falling overhead costs of entering additional markets increases aggregate measures of concentration.

Comparing the results from the model with the observed trends in figure 1, and interpreting the two markets in the model as different product markets within a sector, we see that the only force in the model capable of explaining both increasing sector

---

[23]In figure 8, the sector HHI is not monotone in τ: when trade costs are very low, decreasing τ further tends to decrease the sector HHI. However, the difference between market and sector HHIs is always lower when τ is lower.

[24]Formally, the variance of productivity, $\frac{1}{c}$, is decreasing in k for $k > 2$. When $k \leqslant 2$, the variance of productivity does not exist.

Figure 8: Comparative statics: HHI and welfare
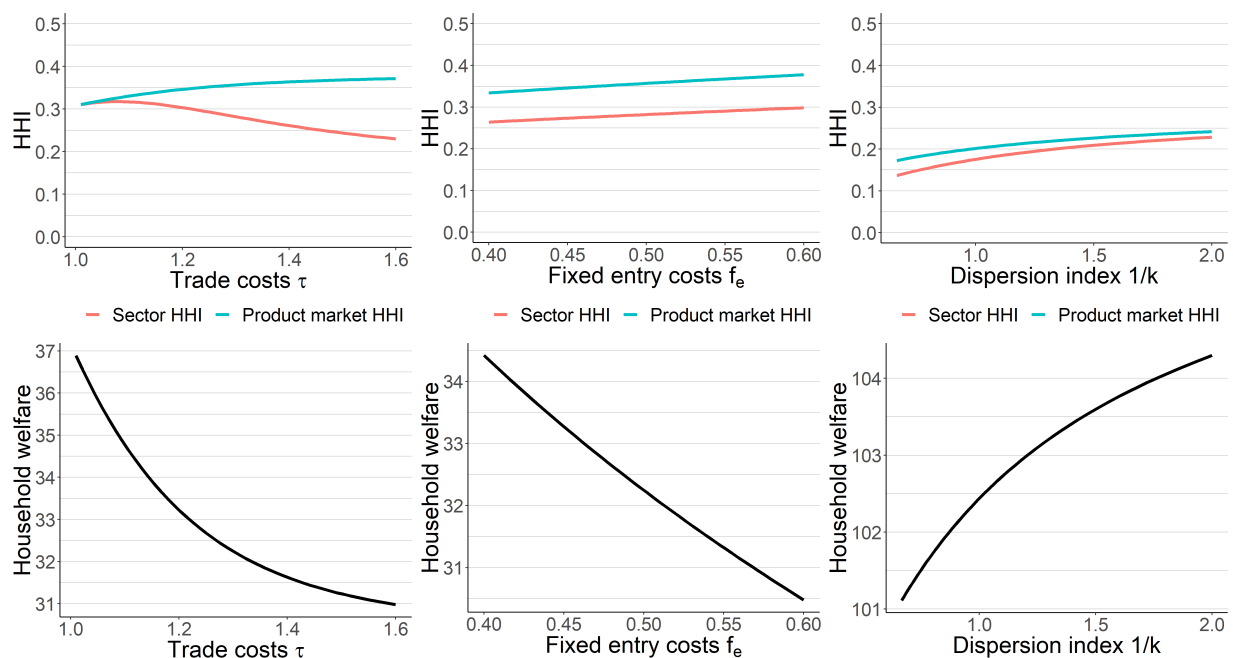
Notes. Comparative statics of product market and sector HHI (top row) and household welfare (bottom row), as we vary trade costs τ (left), fixed entry costs f_E (middle), and firms' productivity dispersion, $\frac{1}{k}$ (right). When $\frac{1}{k}$ is higher, firms' productivity draws are more long-tailed. We vary τ and f_E holding all other parameters fixed. When we vary k, we vary f_E to hold c_D fixed.

concentration and decreasing product market concentration is falling trade costs, τ. Increasing sector concentration can also be explained by increasing fixed costs f_e or by increasing productivity dispersion – the latter explanation being similar to the "superstar firms" hypothesis of (Autor et al., 2017) – and it is possible or even likely that multiple forces are at work. However, both of those forces would also tend to increase product market concentration. To get decreasing product market concentration requires falling trade costs.

Decreasing trade costs could be driven by several factors. It may be that production technologies are becoming more general, perhaps due to increasing use of software in production, making it easier to produce different products in the same plant. It could be that differences in productivity are increasingly driven by general factors such as automated production processes, supply chain efficiency and distribution, and advertising effectiveness, rather than product specific factors. Atalay, Hortaçsu and Syverson (2014) provide evidence that integration into non-competing markets is used to facilitate intangible inputs.

Distinguishing between different drivers of concentration changes is important because they have different implications for consumer welfare. The bottom row of figure 8 graphs welfare with respect to each model primitive. The figure shows that decreases in trade costs and increases in productivity dispersion both lead to higher consumer welfare, while increases in firms' entry costs lead to lower welfare. The model then suggests that the observed trends may be welfare-improving for consumers. Consumers do not care directly about sector-level concentration so, if productive firms are increasingly able to enter multiple product markets, consumers are better off.

Finally, it is a well established fact that corporate profits have been increasing over the last few decades (Barkai (2016), Covarrubias, Gutiérrez and Philippon (2020),De Loecker, Eeckhout and Unger (2020)). Increasing profits could be caused by increasing market power, but our finding of falling concentration in product markets casts some doubt on this explanation. Figure 9 explores how average profits (of surviving firms) in the model vary with respect to model primitives. The figure shows that, according to the model, decreasing trade costs also cause profits to increase. Thus, qualitatively anyway, it is possible that decreasing trade costs could explain both the observed trends in

Figure 9: Comparative statics: Profits
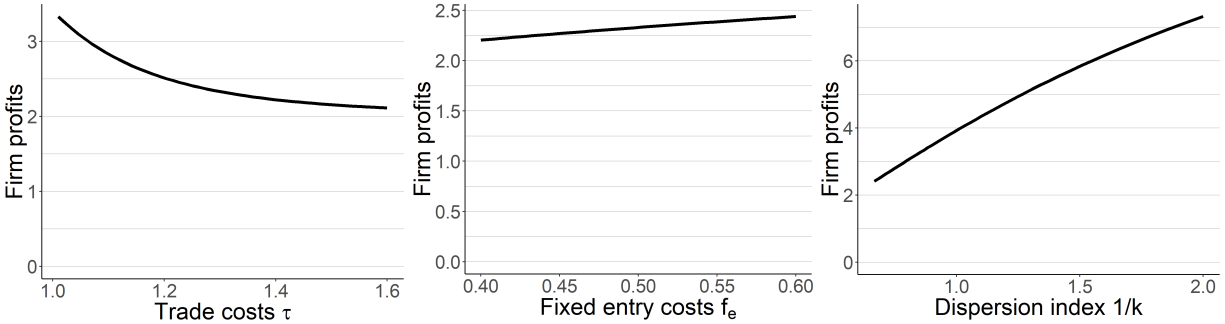
Notes. Comparative statics of firms' expected profits conditional on survival, as we vary trade costs τ (left), fixed entry costs f_E (middle), and firms' productivity dispersion, $\frac{1}{k}$ (right). When $\frac{1}{k}$ is higher, firms' productivity draws are more long-tailed. We vary τ and f_E holding all other parameters fixed. When we vary k, we vary f_E to hold c_D fixed.

concentration and the observed trend in profits. Increasing profits can also be driven by increasing fixed costs (f_E) and increasing productivity dispersion ("superstar firms"), so there may also be multiple forces at work.

5   Conclusion

In conclusion, this paper attempts to measure long term trends in local product market concentration across a wide swath of the U.S. economy, using market definitions that more closely reflect consumption-based economic markets. We find that concentration levels are high in nearly half of the industries covered in our sample, suggesting that market power may be more widespread than previously thought.

We also find that product market concentration has been decreasing over time, particularly in the most concentrated industries. This finding is the opposite of well known results from production data. We do find increasing concentration when product markets are aggregated into sectors, which is consistent with the production data.

An economic model featuring increasing correlation of costs across product markets

explains these two trends. Efficient firms in single product markets enter each others' "home" product markets, thereby increasing aggregate concentration while reducing product level concentration. The model suggests that this process is welfare improving for consumers. While it is possible, or even likely, that multiple forces are at work, this simple model can also explain observed trends in corporate profits.

Trends in product market concentration are an important input to a myriad of literatures in economics. While we have attempted to verify the internal and external validity of our results, confirming them using other consumption-based data sources seems like a pressing issue.

References

Affeldt, Pauline, Tomaso Duso, Klaus Peter Gugler, and Joanna Piechucka. 2021. "Market Concentration in Europe: Evidence from Antitrust Markets."

Aghion, Philippe, Antonin Bergeaud, Timo Boppart, Peter J Klenow, and Huiyu Li. 2019. "A theory of falling growth and rising rents." National Bureau of Economic Research.

Atalay, Enghin, Ali Hortaçsu, and Chad Syverson. 2014. "Vertical integration and input flows." *American Economic Review*, 104(4): 1120–48.

Autor, David, David Dorn, Lawrence F. Katz, Christina Patterson, and John Van Reenen. 2017. "The Fall of the Labor Share and the Rise of Superstar Firms." National Bureau of Economic Research Working Paper 23396. DOI: 10.3386/w23396.

Autor, David, David Dorn, Lawrence F Katz, Christina Patterson, and John Van Reenen. 2020. "The Fall of the Labor Share and the Rise of Superstar Firms*." *The Quarterly Journal of Economics*, 135(2): 645–709.

Barkai, Simcha. 2016. "Declining labor and capital shares." *The Journal of Finance*.

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

Bertrand, Marianne, and Emir Kamenica. 2018. "Coming apart? Cultural distances in the United States over time." National Bureau of Economic Research.

CEA. 2016. "Benefits of competition and indicators of market power." *Issue Brief*.

Covarrubias, Matias, Germán Gutiérrez, and Thomas Philippon. 2020. "From Good to Bad Concentration? US Industries over the past 30 years." *NBER Macroeconomics Annual*, 34(1): 1–46.

Crane, Leland Dod, and Ryan Decker. 2019. "Business Dynamics in the National Establishment Time Series (NETS)."

Crawford, Gregory S, and Ali Yurukoglu. 2012. "The welfare effects of bundling in multichannel television markets." *American Economic Review*, 102(2): 643–85.

Dafny, Leemore, Kate Ho, and Robin S Lee. 2019. "The price effects of cross-market mergers: theory and evidence from the hospital industry." *The RAND Journal of Economics*, 50(2): 286–325.

De Loecker, Jan, Jan Eeckhout, and Gabriel Unger. 2020. "The rise of market power and the macroeconomic implications." *The Quarterly Journal of Economics*.

Doane, Michael, Luke Froeb, Gregory Werden, and David Zimmer. 2018. "Hertz-Dollar Thrifty: Fixing a Merger to Avoid Litigation." *The Antitrust Revolution: Economics, Competition, and Policy .* 7 ed. Oxford University Press.

Economist, The. 2016. "Corporate concentration: the creep of consolidation across America's corporate landscape."

Ganapati, Sharat. 2020. "Growing Oligopolies, Prices, Output, and Productivity." Working Paper.

Gaynor, Martin. 2018. "Examining the Impact of Health Care Consolidation." Statement before the Committee on Energy and Commerce, Oversight and Investigations Subcommittee, U.S. House of Representatives.

Gentzkow, Matthew, and Jesse M Shapiro. 2011. "Ideological segregation online and offline." *The Quarterly Journal of Economics*, 126(4): 1799–1839.

Grieco, Paul, Charles Murry, and Ali Yurukoglu. 2020. "The Evolution of Market Power in the US Auto Industry." Working Paper.

Grullon, Gustavo, Yelena Larkin, and Roni Michaely. 2019. "Are US industries becoming more concentrated?" *Review of Finance*, 23(4): 697–743.

Gutiérrez, Germán, and Thomas Philippon. 2017. "Declining Competition and Investment in the US." National Bureau of Economic Research.

Hsieh, Chang-Tai, and Esteban Rossi-Hansberg. 2019. "The industrial revolution in services." National Bureau of Economic Research.

Keil, Jan. 2017. "The trouble with approximating industry concentration from Compustat." *Journal of Corporate Finance*, 45: 467 – 479.

Melitz, Marc J, and Gianmarco IP Ottaviano. 2008. "Market size, trade, and productivity." *The Review of Economic Studies*, 75(1): 295–316.

Neiman, Brent, and Joseph Vavra. 2018. "The rise in household spending concentration." *Available at SSRN 3137782*.

Nocke, Volker, and Michael D Whinston. 2020. "Concentration screens for horizontal mergers." National Bureau of Economic Research.

Peltzman, Sam. 2014. "Industrial concentration under the rule of reason." *The Journal of Law and Economics*, 57(S3): S101–S120.

Rossi-Hansberg, Esteban, Pierre-Daniel Sarte, and Nicholas Trachter. 2020. "Diverging Trends in National and Local Concentration." *NBER Macroeconomics Annual 2020, volume 35*. University of Chicago Press.

Shapiro, Carl. 2010. "The 2010 horizontal merger guidelines: From hedgehog to fox in forty years." *Antitrust LJ*, 77: 49.

Shapiro, Carl. 2018. "Antitrust in a Time of Populism." *International Journal of Industrial Organization*, 61: 714–748.

Smith, Dominic, and Sergio Ocampo. 2021. "The Evolution of US Retail Concentration."

Wasi, Nada, and Aaron Flaaen. 2015. "Record linkage using Stata: Preprocessing, linking, and reviewing utilities." *The Stata Journal*, 15(3): 672–697.

Internet Appendix

A Data Appendix

A.1 Cleaning Kantar Adspender

From the Kantar Adspender data, we observe brand names, ultimate parents, and product categories. Product category verbal descriptions and codes are available for the years 1994, 2003, and 2017. For the other years, only Kantar's category "codes" are available. The codes appear to be consistent for nearby years, so we impute verbal descriptions for the 1997 data using the 1994 data, and for the 2001 and 2006 data using the 2003 data.

A.2 Merging MRI-Simmons and Kantar Adspender

We use a fuzzy merging algorithm to match brands from the MRI-Simmons data to the Kantar Adspender data. The MRI-Simmons data contain approximately 450 product markets per year, which are relatively stable over time. Kantar Adspender is also divided into around 550 categories, which change somewhat over time. We do the match entirely separately for each year of the dataset.

Data cleaning. We begin by cleaning both datasets, standardizing brand names. We replace accented characters with their closest alphabetic equivalents, remove all non-alphanumeric characters, remove excess whitespace, and lowercase all brands. Additionally, we remove common words such as "and", "any", and "or". Second, from Adspender brands, we remove categorizing words such as "auto" and "corp", which allows longer Adspender brands ("audi auto corp") to match with shorter MRI-Simmons brands ("audi"). Many brands in the Adspender data are very long, including "brand" words followed by "product descriptors", such as "OSCAR DE LA RENTA DRESSES WOMEN". We thus trim brands with many words, by removing either 1 or two words from the end of the brand string; we never trim brands down to less than 3 words.

We manually edit the match, removing around 350 words that are specific enough that they are used for matching by the fuzzy merge algorithm, but are not brand words,

and thus induce bad matches. We also manually delete a few owners and brands which seem to match poorly.

Fuzzy merging. We then merge brand names from the two datasets using a two-step process. We first match MRI-Simmons product markets to Kantar product categories, then run the Stata reclink2 package, created by Wasi and Flaaen (2015), to match MRI brands to Kantar brands. Reclink2 is a fuzzy text merging algorithm, which calculates the distance between strings using a modified bigram algorithm: roughly speaking, this calculates the ratio of the number of common two consecutive letters of the two strings and their average length minus one.

In the first stage, we construct a one-to-many match of MRI-Simmons product markets to Kantar product categories. We first naively fuzzy-merge the full list of MRI-Simmons brands to the full list of Kantar brands. We then check, for each MRI-Simmons product market, the Kantar categories which are matched to the product market most often. We hand-check this merge, adding and subtracting some associations which are not well-captured by the algorithm.

Next, once we have constructed the MRI-Simmons to Kantar category crosswalk, we re-run reclink2, matching brands from MRI-Simmons to Kantar brands within the matched categories. Since the lists of brands to be matched are smaller, false positives are less likely, so we can use a lower match score cutoff.

We use a few more post-processing steps for the merge. In some cases, an MRI-Simmons brand is matched to the same Kantar owner for, for example, 1997 and 2003, but not 2001; this is likely to be a false negative for 2001, so we assign the brand in 2001 to its 1997 and 2003 owner. To improve on the missed matches for brands that have a high market-share, in some cases we manually check brand information using web searches and company websites to assign an owner.

For brands where we are unable to impute an owner using Kantar, we group together brands within the same product category that start with the same first word together; this largely allows us to capture minor products which have the same owner, for example, "Lipton Decaffeinated Iced Tea", "Lipton Iced Tea Mix" and "Lipton Tea & Honey". We then restrict attention to MRI-Simmons product markets for which we are able to impute owners for at least 60% of market share, for all 6 years in our dataset. This reduces the

37

sample from 466 product markets to 337 product markets. In subsection B.3, we also report results for the unbalanced panel, including all 466 product markets.

A.3 Adjusting HHI estimates for finite-sample bias

The HHI is a convex function of market shares. Thus, if we calculate HHIs by plugging in unbiased estimators of market shares, the HHI estimates will tend to be biased upwards. However, this bias can be estimated, allowing us to construct approximately unbiased estimates of the HHI.

Suppose we wish to estimate the market share s_{oj} of owner o in market j. j could represent any level of aggregation, from stategroup-product market, to national sector. Suppose we have some unbiased estimator $\hat{s}_{oj}$ of s_{oj}, that is:

$$E\left[\hat{s}_{oj}\right] = s_{oj}$$

We wish to estimate the HHI in market j, which is:

$$HHI_j = \sum_o s_{oj}^2$$

If we simply estimate the HHI using the sum of squares of estimated market shares, $\hat{s}_{oj}^2$, this will tend to be upwards biased. To see this, note that:

$$E\left[\hat{s}_{oj}^2\right] = \left(E\left[\hat{s}_{oj}\right]\right)^2 + Var\left[\hat{s}_{oj}\right] = s_{oj}^2 + Var\left[\hat{s}_{oj}\right]$$

Hence,

$$E\left(\sum_o \hat{s}_{oj}^2\right) = \sum_o s_{oj}^2 + \sum_o Var\left[\hat{s}_{oj}\right] = HHI_j + \sum_o Var\left[\hat{s}_{oj}\right]$$

Note that, rearranging, we have:

$$HHI_j = E\left(\sum_o \hat{s}_{oj}^2\right) - \sum_o Var\left[\hat{s}_{oj}\right] \tag{14}$$

Hence, we can construct an unbiased estimator for HHI_j by subtracting $\sum_o Var\left[\hat{s}_{oj}\right]$,

the sum of sample variances of the market shares $\hat{s}_{oj}$, from the sample HHI, $\sum_o \hat{s}^2_{oj}$. We calculate these sample variances using a nonparametric bootstrap. In each year of the original survey, we draw 100 samples of users with replacement from the original sample. We calculate market shares at each level of aggregation using these resampled datasets, and take the variance in all market shares over the bootstrap samples. We then use these variances to adjust sample HHIs, using expression (14). All HHI estimates from MRI-Simmons data in the paper include the adjustment in (14).

To illustrate how much the bias correction affects our estimates, figure A.1 compares our HHI estimates from figure 1 to raw HHIs without the adjustment in (14). At the local product market level, the adjustment is fairly large, shifting the estimated median HHI by approximately 130 (out of 10,000). This is because the number of respondents in each local product market is not large – around a few hundred on average – so local product market shares have fairly high sample variances, making the adjustment term in (14) fairly large. In contrast, the adjustment term is essentially negligible at the national product market, local sector, and national sector levels.

Figure A.1: Effect of finite-sample HHI adjustment
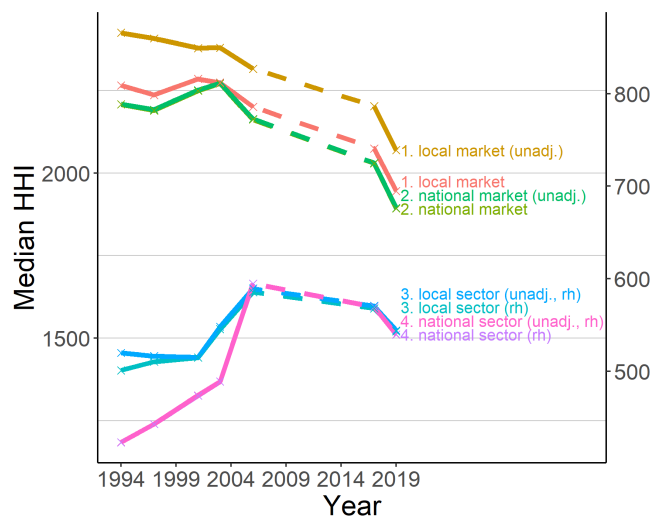
Notes. Effect of the finite-sample HHI bias adjustment, (14), on our HHI estimates. Lines labelled "unadj." are the raw sample HHIs, $\sum_o \hat{s}^2_{oj}$. The other lines are identical to those from Figure 1.

B Robustness checks

Subsection B.1 does a more detailed comparison of our results to the Census data, accounting for changing definitions of NAICS codes over time. Subsection B.2 uses top-2 and top-4 market shares, instead of HHIs, as our measure of concentration. Subsection B.3 uses the entire unbalanced panel of product markets, instead of dropping markets to balance the panel. Subsection B.4 weights sector HHIs by expenditure shares from the Consumer Expenditure Survey. In all three cases, our baseline results are qualitatively and quantitatively unchanged. Finally, subsection B.5 analyzes concentration at two intermediate levels of aggregation between product markets and sectors, and finds that the general trend continues to hold: concentration is increasing over time at higher levels of aggregation, and decreasing at lower levels of aggregation.

B.1 Market coverage – accounting for changing NAICS definitions

In the main text of the paper we show that C4 rises over time at the 6-digit NAICS level for a subsample of Census industries matched to the MRI-Simmons data. One issue with this simple comparison is that NAICS code definitions are changing over time. (Note that this issue is also present in previous work measuring concentration in the Census data. The typical solution in that literature is to drop markets whose definitions changed during the sample period.) To show that NAICS code redefinitions are not driving our results, we conduct the following exercise. We run a simple regression:

$$HHI_{jt} = \mu_t + \gamma_j + \epsilon_{jt} \tag{15}$$

where j indexes NAICS codes, t indexes periods of 5 years, and ϵ_{jt} is an error term that is independent of μ_t and γ_j. If a NAICS code is ever affected by a split or merger, we treat it as a separate NAICS code pre- and post-merger. We are interested in the μ_t coefficients from specification (15). Effectively, (15) is a fixed-effects specification: the time fixed effects μ_t estimate changes in concentration, using only variation within given NAICS codes, over time periods where that code is not affected by code merger events. Specification (15) is a simple way to use all the variation in concentration over time in the

Table A.1: Census HHI over time, fixed effects specification

	(1) All	(2) All unchanged codes	(3) Matched codes	(4) Matched unchanged codes
μ_{2002}	2.550***	2.715***	3.652***	3.671***
	(0.361)	(0.369)	(0.987)	(0.903)
μ_{2007}	3.797***	3.854***	3.928***	4.253***
	(0.374)	(0.379)	(0.978)	(0.942)
μ_{2012}	4.479***	4.789***	3.238**	4.016***
	(0.420)	(0.416)	(1.232)	(1.020)
Constant	32.12***	32.11***	42.52***	42.77***
	(0.301)	(0.310)	(0.809)	(0.760)
N	3429	2737	392	376
R^2	0.930	0.934	0.911	0.932

Standard errors in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Notes. Columns 1 and 2 show the regression results for the entire census sample. Columns 3 and 4 show the results for census codes that we matched to MRI-Simmons product markets. For columns 2 and 4, we also drop any NAICS industry codes that change in the time series.

Census that is not affected by NAICS code redefinitions. The results of specification (15), for various subsets of the data, are shown in table A.1. For all census subsamples, we find that the fixed effects μ_t are increasing uniformly from 1997 to 2012.

B.2   C2 and C4 concentration measures

Figure A.2 replicates figure 1 using two alternative measures of concentration: the sum of the top two owner market shares (C2), and the sum of the top four owner market shares (C4). The trends are similar to using HHI.

Figure A.2: Median C2 and C4 over time, by market definition
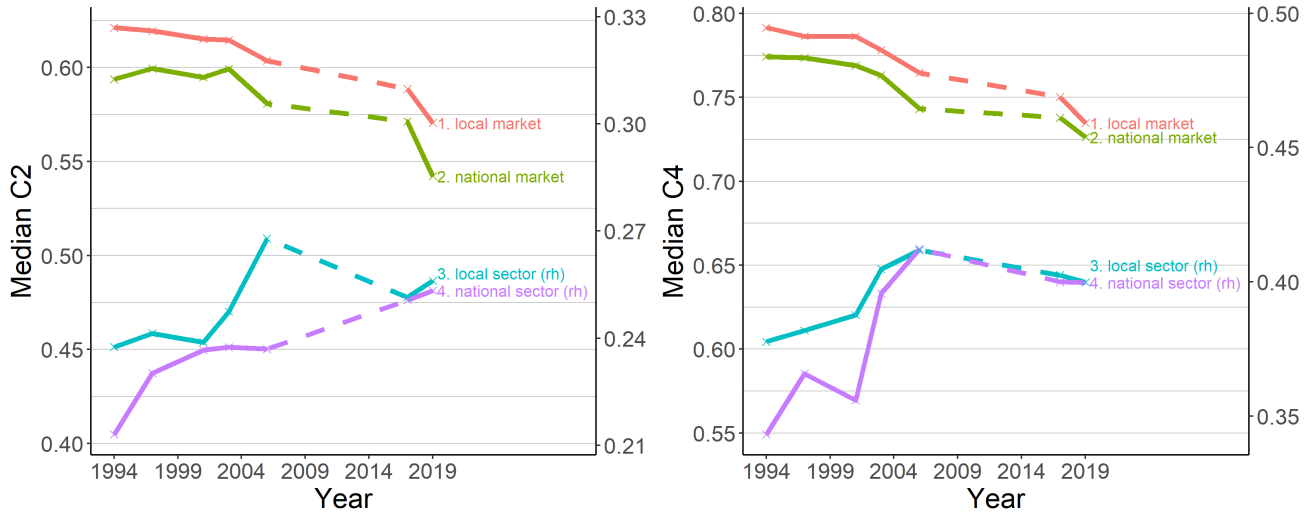
Notes. Local markets are defined as product markets intersected with 29 stategroups. Sectors are defined by aggregating related national product markets. rh indicates right-hand axis.

B.3 Unbalanced panel

Figure A.3 shows the result of figure 1 using all 466 markets we observe in the sample; thus, brand merge rates are lower, and the composition of product markets shifts over time. Nonetheless, the basic pattern that concentration is decreasing at the market level, and somewhat increasing at the sector level, is still present.

B.4 CEX weighting

One concern with our baseline result is that the weights on categories do not reflect consumers' expenditures in these product categories. We attempt to address this by re-weighting results to reflect expenditure shares across broad product categories from the Consumer Expenditure Survey (CEX). To do so, we manually match each item in CEX to a MRI-Simmons sector. For example, we match "Cereals and bakery products" to "HoProdFood" and "Nonalcoholic beverages" to "Beverages". This is a many-to-1 match: CEX items are available in fine subitems, so in general there are many CEX items within

Figure A.3: Median HHI at different market levels over time, unbalanced panel
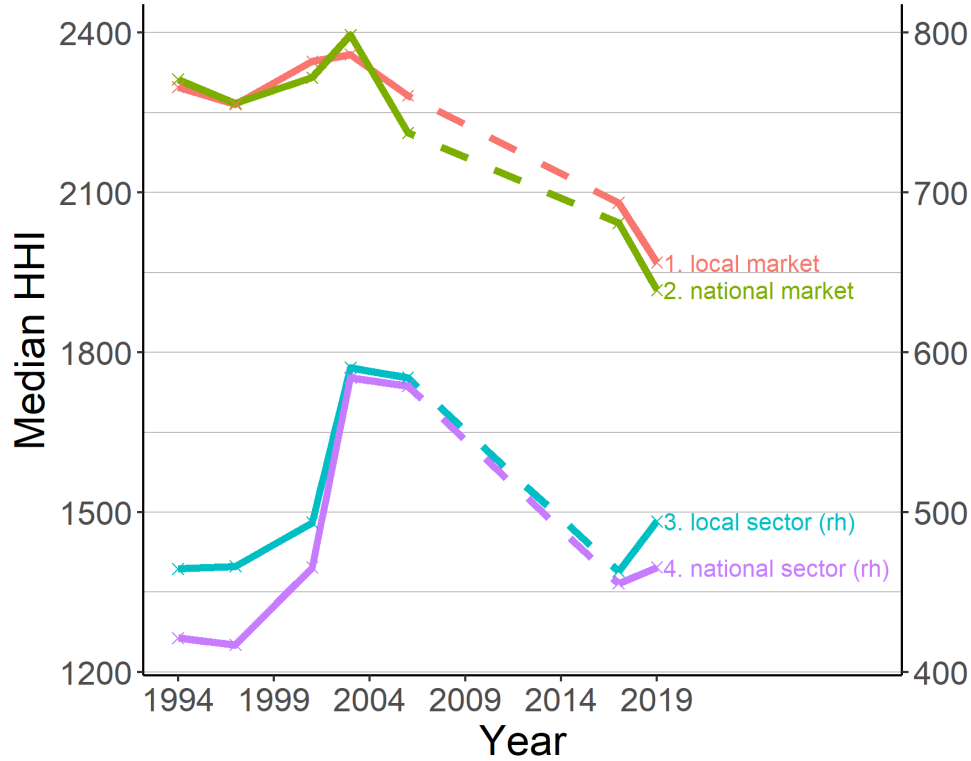
Notes. Median HHI over time, at the state-market, market, state-sector, and sector levels, for the unbalanced panel dataset. "rh" indicates right-hand axis.

each MRI-Simmons sector. We then aggregate expenditure shares from the 2018 CEX table to MRI-Simmons sector level, and we use these expenditure shares to weight sectors in our baseline result.

The results are shown in figure A.4. The results are qualitatively quite similar to results in figure 1 in the main text. Local and national market-level concentration both decline substantially. Local sector-level concentration increases slightly, while national sector-level concentration rises and then falls, so it is essentially flat over the entire time horizon.

Figure A.4: CEX Weighting
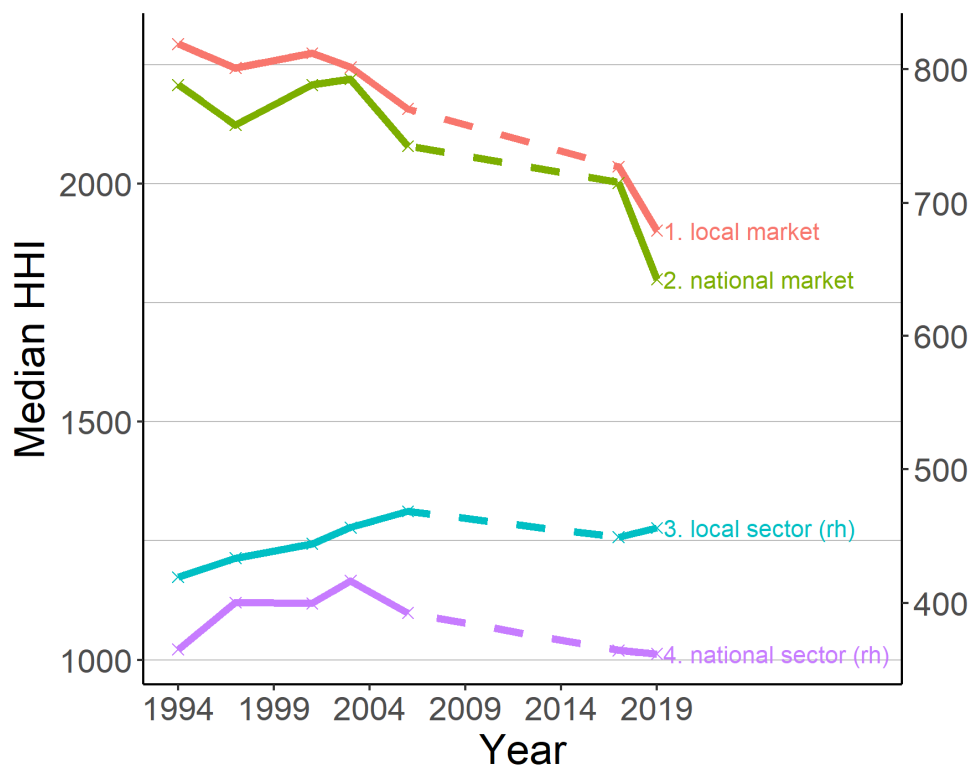
Notes. Equivalent of figure 1, where we re-weight sectors, so they reflect expenditure shares from the CEX. Each line shows the CEX-expenditure-weighted median of HHIs. "rh" indicates right-hand axis.

B.5 Alternative levels of market aggregation

One weakness of our main data is that we only have two market definitions: markets, and sectors. However, the Kantar dataset, for 2017, has multiple levels of market aggregation: "majors" and "industries", which are somewhat lower-level than MRI-Simmons sectors. While we only have these aggregation variables for a single year of the Kantar data, if we hold fixed the mapping from markets to majors and industries over time, we can use this to analyze concentration at different levels of market aggregation. That is, we first match MRI-Simmons markets to fine Kantar product categories; we then impute Kantar "majors" and "industries" using the 2017 Kantar definitions. This gives us two more levels of aggregation for analyzing concentration: in the balanced panel, we have

337 MRI-Simmons product markets, 120 Kantar majors, 46 Kantar industries, and 17 MRI-Simmons sectors.

In figure A.5, we show how concentration varies at each of these levels of aggregation. MRI-Simmons markets are the finest level of aggregation, followed by Kantar majors, Kantar industries, and MRI-Simmons sectors. The "divergence" trend is relatively uniform. Concentration is decreasing over time at the MRI-Simmons market level, roughly flat at the Kantar major and industry levels, and increasing over time at the MRI-Simmons sector level.

Figure A.5: Median HHI including Kantar major and industry levels
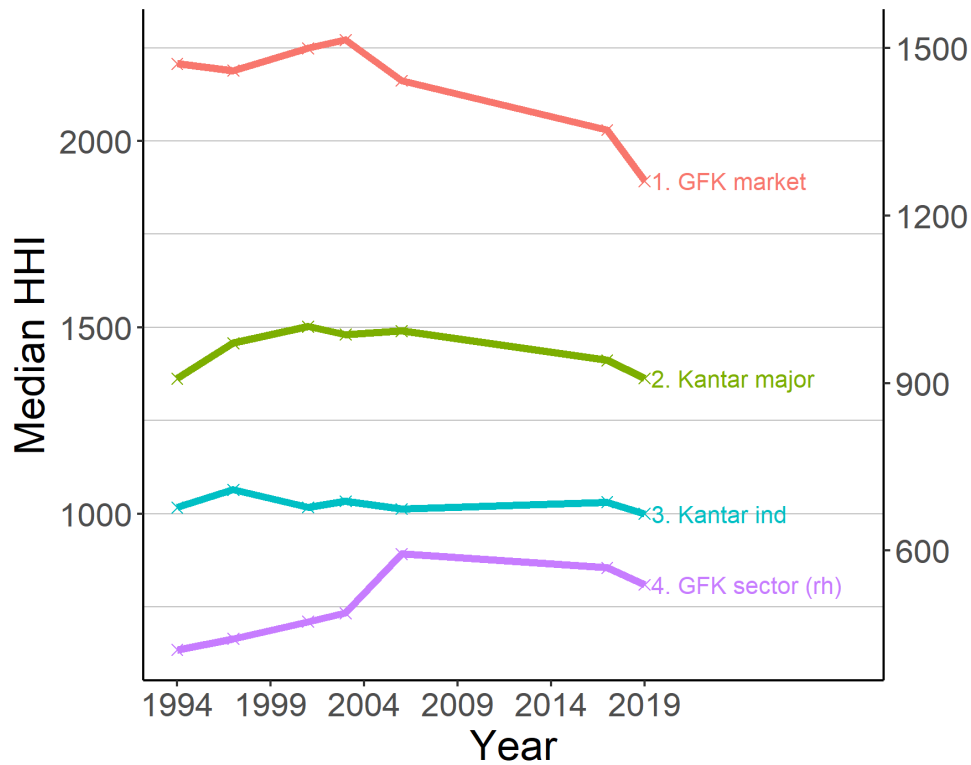
Notes. Equivalent of figure 1, including Kantar major (green) and industry (blue) levels. MRI-Simmons product markets (red) and MRI-Simmons sectors (purple) are identical to figure 1. rh indicates right-hand axis.

B.6   HHI by sector and aggregation level

To complement figure 4, local product market HHI over time by sector, figure A.6 shows national product market, local sector HHI, and national sector HHI over time by sector, in addition to local product market HHI.

C   Proofs for section 4

C.1   Proof of proposition 1

The proof essentially follows directly from Melitz and Ottaviano. Expression (4) for the cost cutoff is (29) in Melitz and Ottaviano, with $L^l = 1$, and using that $\rho \equiv \frac{1}{\tau}$. The technology parameter, ϕ, is defined on page 304 of Melitz and Ottaviano. Expression (5) is exactly (16) of Melitz and Ottaviano, which applies in the both the single-market and the two-market case. Expressions (6) and (7) are expression (20) of Melitz and Ottaviano, with $L^l = 1$ and $L^h = 1$ and symmetric trade costs.

C.2   Proof of proposition 2

We prove proposition 2 in a few steps. First, claims 1 and 2 characterize two properties of HHIs in the continuous case.

C.2.1   Renormalization properties of HHIs

Claim 1 illustrates how the continuous HHI is analogous to the familiar discrete HHI. If all firms produce identical quantities, the HHI is simply $\frac{1}{N}$, so it is decreasing in the measure of firms which produce. Holding fixed N, the HHI is higher when the dispersion of firms' quantities, measured by the coefficient of variation of $q(c)$, is larger. Expression (16) also holds for discrete HHIs. Thus, the continuous HHI behaves in a way that is qualitatively quite similar to its discrete analog.

Claim 1. Let $H(c)$ represent the distribution of costs among firms producing positive quantities, and let N represent the number of firms producing in a given market. We

have:

$$HHI_j = \frac{1}{N}\left(1 + \left(\frac{SD\left(q\left(c\right)\right)}{E\left[q\left(c\right)\right]}\right)^2\right) \tag{16}$$

Proof. It will be useful to work with, rather than the measure space i, the distribution of costs among firms producing strictly positive amounts in each markets. Suppose there are N firms – domestic firms and exporters – which produce positive quantities in each market. We can disregard firms which produce zero quantity, as they contribute nothing to the HHI integral.

Defining the probability distribution $H\left(c\right)$ of costs of firms who produce positive quantities, and $G\left(c\right)$ the distribution of costs among firms that produce, we have, for any function $f\left(\cdot\right)$, that:

$$\int f\left(c\left(i\right)\right)di = N\int_0^\infty f\left(c\right)dG\left(c\right) \tag{17}$$

Hence, we can write the HHI in terms of the firm cost distribution as:

$$HHI_j = \int_0^N \left(s_j\left(c\left(i\right)\right)\right)^2 dG\left(c\right) = \frac{\int_0^N \left(q_j\left(c\left(i\right)\right)\right)^2 di}{\left[\int_0^N q_j\left(c\left(i\right)\right)di\right]^2}$$

$$= \frac{N\int_0^{c_M}\left(q_j\left(c\right)\right)^2 dG\left(c\right)}{\left(N\int_0^{c_M}q_j\left(c\right)dG\left(c\right)\right)^2} = \frac{1}{N}\frac{\int_0^{c_M}\left(q_j\left(c\right)\right)^2 dG\left(c\right)}{\left(\int_0^{c_M}q_j\left(c\right)dG\left(c\right)\right)^2}$$

$$HHI_j = \frac{1}{N}\frac{E\left[\left(q_j\left(c\right)\right)^2\right]}{\left(E\left[q_j\left(c\right)\right]\right)^2} \tag{18}$$

Expression (18) expresses HHI_j in terms of the first and second moments of the quantity distribution among producing firms, and the total number of producing firms N. We can further rearrange (18) to:

$$= \frac{1}{N}\frac{\left(E\left[q\left(c\right)\right]\right)^2 + Var\left(q\left(c\right)\right)}{\left(E\left[q\left(c\right)\right]\right)^2} = \frac{1}{N}\left(1 + \frac{Var\left(q\left(c\right)\right)}{\left(E\left[q\left(c\right)\right]\right)^2}\right)$$

$$= \frac{1}{N}\left(1 + \left(\frac{SD\left(q\left(c\right)\right)}{E\left[q\left(c\right)\right]}\right)^2\right)$$

as desired. $\qquad\square$

C.2.2 The SVI

First, we show that the difference between the HHI of an aggregated market, and the weighted average of HHI's in submarkets, is related to the variance of market share vectors in submarkets. In particular, consider the following quantity.

Definition 2. Define the continuous SVI (submarket variation index) as:

$$SVI \equiv \int_0^N \sum_j \frac{Q_j}{\sum_j Q_j} \left(s_j(i) - s(i)\right)^2 di \tag{19}$$

Intuitively, equation (19) calculates, for each firm, the Q_j-weighted average of deviations of the firm's market shares in each market from the firm's sector average market share, then we take the integral of this average across all firms. Claim 2 states that the submarket variation index is equal to the W_j-weighted average of product market HHIs, minus the sector HHI. The SVI is thus a simple measure of the divergence between product market and sector HHIs.

Claim 2. The quantities SVI, HHI_j, and HHI satisfy:

$$SVI = \left[\sum_j \frac{Q_j}{\sum_j Q_j} HHI_j\right] - HHI \tag{20}$$

Proof. Take the RHS of (19), and substitute for $s(i)$ using (9), to get:

$$= \int_0^N \sum_j \frac{Q_j}{\sum_j Q_j} \left(s_j(i) - \frac{\sum_j Q_j s_j(i)}{\sum_j Q_j}\right)^2 di$$

Expanding, we have

$$\int_i \sum_j \left(\frac{Q_j}{\sum_j Q_j} (s_j(i))^2 - 2\frac{Q_j}{\sum_j Q_j} (s_j(i)) \left(\frac{\sum_j Q_j s_j(i)}{\sum_j Q_j}\right) + \frac{Q_j}{\sum_j Q_j} \left(\frac{\sum_j Q_j s_j(i)}{\sum_j Q_j}\right)^2\right) di \tag{21}$$

48

This simplifies to:

$$\left[\int_i \left(\sum_j \frac{Q_j}{\sum_j Q_j} (s_j(i))^2 \right) di \right] - \int_i \left(\frac{\sum_j Q_j s_j(i)}{\sum_j Q_j} \right)^2 di$$

Now, recalling that the definition of $s(i)$ in (9), we have:

$$\int_i \left(\frac{\sum_j Q_j s_j(i)}{\sum_j Q_j} \right)^2 di = \int_i (s(i))^2 di = \text{HHI}$$

And,

$$\int_i \left(\sum_j \frac{Q_j}{\sum_j Q_j} (s_j(i))^2 \right) di = \sum_j \frac{Q_j}{\sum_j Q_j} \int_i (s_j(i))^2 di = \sum_j \frac{Q_j}{\sum_j Q_j} \text{HHI}_j$$

This proves (20), and thus claim 2. □

C.2.3 Computing product market HHIs

We know that the distribution of costs (inclusive of trade costs) among producing firms, both domestic firms and exporters, is such that $\frac{1}{c}$ is Pareto distributed, with shape parameter k and minimum $\frac{1}{c_D}$. Also, the quantity $q_j(c)$ is proportional to the difference

$$c_D - c$$

Expression (18) implies that HHI_j is invariant to scaling $q_j(c)$; that is, if we multiply all firms' quantities by some constant k, HHI_j is unchanged. We will thus scale firms' quantities such that a firm with cost 0 has quantity 1. We can also change variables to work with indexed costs

$$x \equiv \frac{c}{c_D}$$

if $\frac{1}{c}$ is Pareto distributed with minimum $\frac{1}{c_D}$, then $\omega = \frac{1}{x}$ is Pareto distributed, with minimum 1 and shape parameter k. Thus, define the rescaled quantity $\tilde{q}(\omega)$; a firm with

49

productivity ω has rescaled quantity:

$$\tilde{q}(\omega) = 1 - \frac{1}{\omega}$$

Thus, we have:

$$E\left[\left(1 - \frac{1}{\omega}\right)^2\right] = \frac{2}{(1+k)(2+k)}, \quad E\left[1 - \frac{1}{\omega}\right] = \frac{1}{1+k}$$

thus, plugging into (18), we have:

$$HHI_j = \frac{2}{N}\frac{1+k}{2+k} \tag{22}$$

Where N is the number of firms producing in a given market, including both domestic firms and exporters. This proves (12) of proposition 2.

Now, let M be the number of entering firms with cost below c_D in one market; this is:

$$N = M\left(1 + \frac{G_c\left(\frac{c_D}{\tau}\right)}{G_c(c_D)}\right)$$

In renormalized space, we have:

$$\left(1 + \frac{G_c\left(\frac{c_D}{\tau}\right)}{G_c(c_D)}\right) = \left(1 + G_c\left(\frac{1}{\tau}\right)\right) = 1 + (1 - G_\omega(\tau)) = 1 + \frac{1}{\tau^k}$$

Hence, the product market HHI is:

$$HHI_j = \frac{1}{M\left(1 + \frac{1}{\tau^k}\right)}\frac{2+2k}{2+k} \tag{23}$$

where M is the measure of domestic firms that produce positive quantities in each market. Note that $M < N$, since N counts both domestic and exporting firms. In particular, since the N firms active in a given market consist of importers as well as a fraction $\frac{1}{\tau^k}$ of

exporters; hence, we have:

$$N = M\left(1 + \frac{1}{\tau^k}\right)$$
(24)

Plugging (24) into (23), we get (12) of Proposition 2. To show that product market HHIs are always increasing in τ, note that from (4) of Proposition 1, an increase in τ causes to c_D to increase. From (5), and increase in c_D causes N to decrease. From (12), a decrease in N causes HHI_j to increase.

C.2.4 Computing the SVI

First, the SVI expression simplifies in the two market symmetric case. This is because:

$$SVI \equiv \int_0^{2M} \sum_j \frac{Q_j}{\sum_j Q_j} \left(s_j(i) - s(i)\right)^2 di = \int_0^{2M} 2\frac{1}{2}\left(s_1(i) - \frac{s_1(i) + s_2(i)}{2}\right)^2 di$$

$$SVI = \frac{1}{4}\int_0^{2M} \left(s_1(i) - s_2(i)\right)^2 di$$

Applying (17) from the proof of claim 1, we can express this as:

$$SVI = \frac{1}{4}\int_0^{c_D}\left(\frac{q_1(c)}{\int q_1(c)\left[M\left(1+\frac{1}{\tau^k}\right)\right]dG(c)} - \frac{q_2(c)}{\int q_2(c)\left[M\left(1+\frac{1}{\tau^k}\right)\right]dG(c)}\right)^2 2M dG(c)$$

Now, by symmetry, the distributions of q_1 and q_2 are identical, so we can write this as:

$$SVI = \frac{1}{2M\left(1+\frac{1}{\tau^k}\right)^2 (E[q_1])^2}\int_0^{c_D}\left(q_1(c) - q_2(c)\right)^2 dG(c)$$
(25)

Analogous to (18), expression (25) expresses SVI in terms of moments of the joint distribution of q_1 and q_2. It is also invariant to scaling, so we will again scale firms' quantities to have maximum 1, and index firms by their cost index, $\omega \equiv \frac{c_D}{c}$, to get:

$$SVI = \frac{1}{2M\left(1+\frac{1}{\tau^k}\right)^2 (E[q_1])^2}\int_0^\omega\left(\tilde{q}_1(\omega) - \tilde{q}_2(\omega)\right)^2 dG_\omega(\omega)$$
(26)

51

where ω is Pareto distributed, with shape k and minimum 1. Firms export only if $c < \frac{c_D}{\tau}$, which is equivalent to $\omega > \tau$.

Now, the value of the SVI integrand,

$$(\tilde{q}_1(\omega) - \tilde{q}_2(\omega))^2$$

differs depending on whether a firm is an exporter or not. For firms in region 1 that do not export, we have

$$\tilde{q}_1(\omega) = 1 - \frac{1}{\omega}, \ \tilde{q}_2(\omega) = 0$$

For firms that do export, we have

$$\tilde{q}_1(\omega) = 1 - \frac{1}{\omega}, \ \tilde{q}_2(\omega) = 1 - \frac{\tau}{\omega}$$

We can thus write (26) as:

$$SVI = \frac{1}{2M\left(1 + \frac{1}{\tau^k}\right)^2 (E[q_1])^2}$$
$$\left[\int_1^\tau \left(1 - \frac{1}{\omega}\right)^2 dG_\omega(\omega) + \int_\tau^\infty \left(\left(1 - \frac{1}{\omega}\right) - \left(1 - \frac{\tau}{\omega}\right)\right)^2 dG_\omega(\omega)\right]$$

We can write this as:

$$SVI = \frac{1}{2M\left(1 + \frac{1}{\tau^k}\right)^2 (E[q_1])^2}$$
$$\left[\int_\tau^\infty \left(\left(1 - \frac{1}{\omega}\right) - \left(1 - \frac{\tau}{\omega}\right)\right)^2 dG_\omega(x) + \int_1^\infty \left(1 - \frac{1}{\omega}\right)^2 dG_\omega(\omega) - \int_\tau^\infty \left(1 - \frac{1}{\omega}\right)^2 dG_\omega(\omega)\right]$$

$$SVI = \frac{1}{2M \left(1 + \frac{1}{\tau^k}\right)^2 (E\,[q_1])^2}$$

$$\left[\int_1^\infty \left(1 - \frac{1}{\omega}\right)^2 dG_\omega(\omega) - \int_\tau^\infty \left(1 - \frac{\tau}{\omega}\right) \left(2\left(1 - \frac{1}{\omega}\right) - \left(1 - \frac{\tau}{\omega}\right)\right) dG_\omega(x)\right]$$

Now, by properties of the Pareto distribution, we have:

$$\int_1^\infty \left(1 - \frac{1}{\omega}\right) dG_\omega(\omega) = \frac{1}{k+1}$$

$$\int_1^\infty \left(1 - \frac{1}{\omega}\right)^2 dG_\omega(\omega) = \frac{2}{(k+1)\,(k+2)}$$

hence,

$$SVI = \frac{(k+1)}{M \left(1 + \frac{1}{\tau^k}\right)^2 (k+2)} - \frac{(k+1)^2}{2M \left(1 + \frac{1}{\tau^k}\right)^2} \int_\tau^\infty \left(1 - \frac{\tau}{\omega}\right) \left(2\left(1 - \frac{1}{\omega}\right) - \left(1 - \frac{\tau}{\omega}\right)\right) dG_\omega(x)$$

(27)

Factoring, and using (24), we get:

$$SVI = \frac{k+1}{N \left(1 + \frac{1}{\tau^k}\right) (k+2)} \left[1 - \frac{1}{2}(k+1)\,(k+2) \int_\tau^\infty \left(1 - \frac{\tau}{\omega}\right) \left(2\left(1 - \frac{1}{\omega}\right) - \left(1 - \frac{\tau}{\omega}\right)\right) dG_\omega(x)\right]$$

(28)

Applying claim 2, we get (13) of proposition 2.

To prove that SVI is increasing in τ, we have:

$$\frac{\partial}{\partial \tau} \left(1 - \frac{\tau}{\omega}\right) \left(2\left(1 - \frac{1}{\omega}\right) - \left(1 - \frac{\tau}{\omega}\right)\right) = \frac{2 - 2\tau}{\omega^2}$$

which is weakly negative for $\tau \geqslant 1$. Hence, as τ increases, the (negative) integral term in (28),

$$\frac{1}{2}(k+1)\,(k+2) \int_\tau^\infty \left(1 - \frac{\tau}{\omega}\right) \left(2\left(1 - \frac{1}{\omega}\right) - \left(1 - \frac{\tau}{\omega}\right)\right) dG_\omega(x)$$

decreases in magnitude. Also, the coefficient term

$$\frac{k+1}{N\left(1+\frac{1}{\tau^k}\right)(k+2)}$$

decreases in magnitude. Thus, the SVI is increasing in τ. This completes the proof of proposition 2.

D   Data coverage

Table A.2 shows the names of each MRI-Simmons product market in the balanced panel. Product market and sector names are defined by MRI-Simmons.

Table A.2: Balanced panel product market coverage

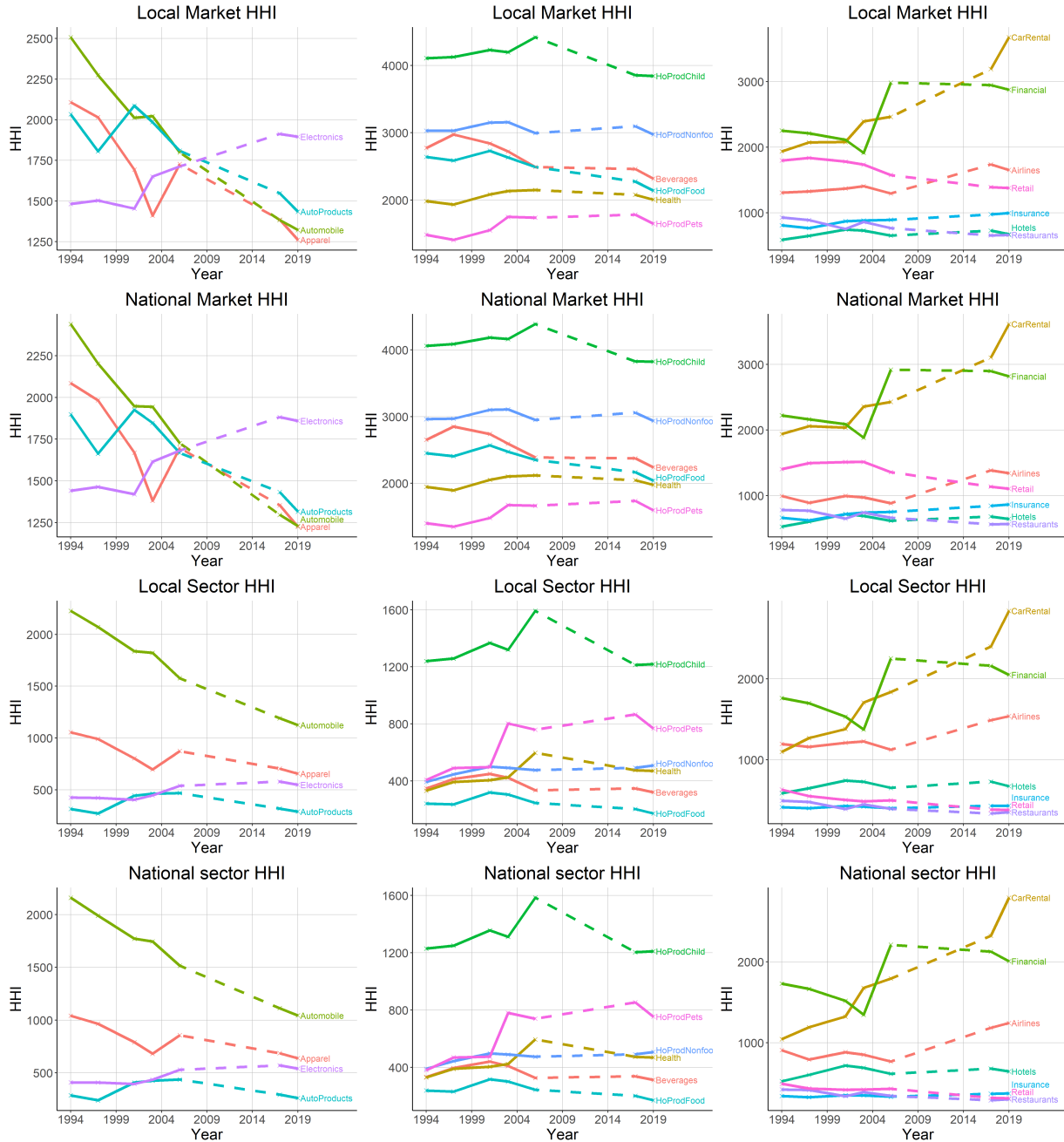| Product market	Sector	Product market	Sector	Product market	Sector
DomesticTravelAirlinesused	Airlines	AdhesiveBandages	Health	InfantCereal	HoProdChild
ForeignTravelAirlinesUsed	Airlines	AftershaveLotionCologneForMen	Health	PainRelieversFeverReducersForChildren	HoProdChild
AthleticShoesBrandsBought	Apparel	AthletesFootFootCareProducts	Health	PreMoistenedBabyWipes	HoProdChild
WomensLingerieUndergarments	Apparel	BathShowerAdditivesWomen	Health	TeethingRemedies	HoProdChild
AirFilters	AutoProducts	BleachAndDepilatories	Health	VitaminsForChildren	HoProdChild
AntiFreezeCoolant	AutoProducts	BlusherWomen	Health	AmericanPasteurizedProcessedCheese	HoProdFood
CarBatteriesBrands	AutoProducts	BodyPowder	Health	ArtificialSweeteners	HoProdFood
CarWaxPolish	AutoProducts	ColdSinusAndAllergyRemediesNonprescr	Health	BaconAndBreakfastStrips	HoProdFood
Gasoline	AutoProducts	ComplexionCareProducts	Health	BakingChips	HoProdFood
GasolineAdditives	AutoProducts	CondomsBought	Health	BakingCoconut	HoProdFood
LeatherAndVinylProtectants	AutoProducts	ContactLensCleaningWettingSolutions	Health	BakingMixesExcludingCakeMixes	HoProdFood
MotorOil	AutoProducts	CoughSyrupNonprescription	Health	BakingPowderAndSoda	HoProdFood
MotorOilAdditives	AutoProducts	DentalFloss	Health	BarBakingChocolate	HoProdFood
MufflersBrands	AutoProducts	DentureAdhesivesFixatives	Health	BottledBarbecueSeasoningSauces	HoProdFood
OilFilters	AutoProducts	DentureCleaners	Health	BoxedChocolates	HoProdFood
ShockAbsorbersStruts	AutoProducts	DeodorantsAndAntiperspirants	Health	Bread	HoProdFood
SparkPlugs	AutoProducts	DisposableRazors	Health	BreadCrumbsCoatingMixes	HoProdFood
Tires	AutoProducts	ElectricShavers	Health	BreakfastCerealGranolaBars	HoProdFood
WindshieldWipers	AutoProducts	EyeLinerWomen	Health	BreakfastCerealsCold	HoProdFood
AutomobilesAndOtherVehiclesManufacturer	Automobile	EyeShadowWomen	Health	BreakfastCerealsHot	HoProdFood
MotorcyclesMake	Automobile	EyeWashAndDrops	Health	BrownieCookieMixes	HoProdFood
BottledWaterSeltzer	Beverages	FacialMoisturizersWomen	Health	Butter	HoProdFood
Bourbon	Beverages	FeminineHygieneDeodorantCleansingProducts	Health	CandyRegularOrKingSize	HoProdFood
Brandy	Beverages	FeminineMedicatedProductsWomen	Health	CannedBeansWithSauce	HoProdFood
CanadianWhisky	Beverages	FoundationMakeUpWomen	Health	CannedChicken	HoProdFood
ChampagneSparklingWines	Beverages	HairColoringProductsForUseAtHome	Health	CannedOrJarredFruit	HoProdFood
Cognac	Beverages	HairConditionersForUseAtHome	Health	CannedOrJarredSoup	HoProdFood
CordialsLiqueurs	Beverages	HairConditioningTreatmentForUseAtHome	Health	CannedOrJarredSpaghettiMacaroni	HoProdFood
DietColaDrinks	Beverages	HairSpraysForUseAtHome	Health	CannedOrJarredVegetables	HoProdFood
DomesticDinnerTableWines	Beverages	HairStylingGelsLotions	Health	CannedStews	HoProdFood
EnergyDrinks	Beverages	HairTonicOrDressingMen	Health	CannedTomatoes	HoProdFood
EvaporatedCondensedMilk	Beverages	HandBodyCreamLotionOrOil	Health	ChewingGum	HoProdFood
FlavoredAlcoholicBeveragesCoolers	Beverages	HeadacheRemediesAndPainRelieversNonprescr	Health	ChickenTurkeyFreshOrFrozen	HoProdFood
FlavoredInstantCoffee	Beverages	HemorrhoidRemedies	Health	Chili	HoProdFood
Gin	Beverages	HomePermanentsWomen	Health	CocoaPowder	HoProdFood
GrapefruitJuice	Beverages	Laxatives	Health	ColdCuts	HoProdFood
ImportedBeer	Beverages	LipCare	Health	CookedHams	HoProdFood
ImportedDinnerTableWines	Beverages	LipstickLipGlossWomen	Health	CookiesReadyToEat	HoProdFood
LowCalorieDomesticBeer	Beverages	LiquidSoapsHandSanitizers	Health	CottageCheese	HoProdFood
MaltLiquor	Beverages	LooseFacePowderWomen	Health	Crackers	HoProdFood
OtherDietSoftDrinksNotColas	Beverages	MascaraWomen	Health	CreamCheese	HoProdFood
OtherFruitJuicesDrinks	Beverages	MealSupplements	Health	DinnerMixesAndKits	HoProdFood
OtherRegularCarbonatedSoftDrinks	Beverages	MedicatedSkinOintments	Health	DrinkAdditivesHotCocoaAddMilkOrWater	HoProdFood
PopularDomesticBeer	Beverages	MedicatedThroatRemedies	Health	DryCakeMixes	HoProdFood
PortSherryDessertWines	Beverages	Mouthwash	Health	DrySoupBouillon	HoProdFood
PowderedFruitSoftDrinks	Beverages	NailCareProductsPolishWomen	Health	EggAlternatives	HoProdFood
PremiumDomesticBeer	Beverages	NailPolishRemoverWomen	Health	EnglishMuffins	HoProdFood
PreparedMixedDrinksWithoutLiquor	Beverages	PainRelievingRubsLiquidsNonprescription	Health	Extracts	HoProdFood
ReadyToDrinkIcedTea	Beverages	PersonalCareSoapsBar	Health	FishSeafoodFreshOrFrozen	HoProdFood
RegularColaDrinksNotDiet	Beverages	RazorBlades	Health	FlavoredSeasonedRice	HoProdFood
RegularDomesticBeer	Beverages	SanitaryNapkinsAndPantilinersWomen	Health	FlourCornmeal	HoProdFood
RegularTea	Beverages	ShampooForUseAtHome	Health	FrankfurtersWieners	HoProdFood
Rum	Beverages	ShavingCreamsOrGels	Health	Frostings	HoProdFood
RyeOrBlendedWhiskey	Beverages	SleepingTabletsNonprescription	Health	FrozenBreadedChicken	HoProdFood
ScotchWhisky	Beverages	SuntanSunscreenProducts	Health	FrozenBreakfasts	HoProdFood
SuperPremiumDomesticBeer	Beverages	TamponsWomen	Health	FrozenCompleteDinners	HoProdFood
Tequila	Beverages	ToothacheGumCankerSoreRemedies	Health	FrozenDesserts	HoProdFood
TomatoAndVegetableJuices	Beverages	Toothbrushes	Health	FrozenHotSnacks	HoProdFood
Vermouth	Beverages	Toothpaste	Health	FrozenMainCourses	HoProdFood
Vodka	Beverages	Toothpolish	Health	FrozenOrangeJuice	HoProdFood
CarRentalBusinessUse	CarRental	VitaminAndMineralSupplements	Health	FrozenPancakesFrenchToast	HoProdFood
CarRentalPersonalUse	CarRental	WartRemovers	Health	FrozenPizza	HoProdFood
TruckTrailerRentalCompanies	CarRental	BabyBathWashAndSoap	HoProdChild	FrozenRefrigeratedPotatoProducts	HoProdFood
Batteries	Electronics	BabyFoods	HoProdChild	FrozenVegetables	HoProdFood
CamerasCamcordersBrands	Electronics	BabyLotion	HoProdChild	FrozenWaffles	HoProdFood
DVDBluRayPlayersBrands	Electronics	BabyNursers	HoProdChild	FrozenYogurt	HoProdFood
PersonalComputers	Electronics	BabyOil	HoProdChild	GelatinAndGelatinDesserts	HoProdFood
TelevisionSetsBrands	Electronics	BabyOintments	HoProdChild	GratedCheese	HoProdFood
CreditCards	Financial	BabyPowder	HoProdChild	GravySauceMixesAndCookingSauces	HoProdFood
InvestmentActivityBrokerageFirms	Financial	BabyShampoo	HoProdChild		
RealEstateWhichAgentUsed	Financial	ChildrensColdTabletsLiquids	HoProdChild		
		ChildrensCoughSyrup	HoProdChild		
		CottonSwabs	HoProdChild		
		DisposableDiapers	HoProdChild		

Notes. MRI-Simmons product market names and sector names in our data.

Table A.2 (continued): Balanced panel product market coverage

Product market	Sector	Product market	Sector
GroundCoffee	HoProdFood	WaffleOrPancakeMix	HoProdFood
HardRollCandy	HoProdFood	WhippedTopping	HoProdFood
Honey	HoProdFood	WholeCoffeeBeans	HoProdFood
IceCreamBarsSandwichesBonBons	HoProdFood	Yeast	HoProdFood
IceCreamIceMilkSherbet	HoProdFood	Yogurt	HoProdFood
InstantBreakfast	HoProdFood	AirFreshenersCarpetRoomDeodorizers	HoProdNonfood
InstantIcedTeaMix	HoProdFood	AluminumFoil	HoProdNonfood
InstantOrFreezeDriedCoffee	HoProdFood	AutomaticDishwasherDetergent	HoProdNonfood
JamsJellies	HoProdFood	Bleach	HoProdNonfood
KetchupCatsup	HoProdFood	CarpetAndRugCleaners	HoProdNonfood
LunchCombinationsKits	HoProdFood	Charcoal	HoProdNonfood
Margarine	HoProdFood	CharcoalLighterFluid	HoProdNonfood
MayonnaiseAndMayonnaiseTypeSaladDressing	HoProdFood	ChewingAndSmokelessTobacco	HoProdNonfood
MeatSnacks	HoProdFood	Cigars	HoProdNonfood
MexicanFoods	HoProdFood	DishwashingLiquid	HoProdNonfood
Mints	HoProdFood	DisposableCups	HoProdNonfood
Mustard	HoProdFood	DisposablePlates	HoProdNonfood
NaturalOrImportedCheese	HoProdFood	DrainCleaners	HoProdNonfood
Nectars	HoProdFood	FabricSofteners	HoProdNonfood
NonDairyCreamSubstitutes	HoProdFood	FacialTissues	HoProdNonfood
Nuts	HoProdFood	Firelogs	HoProdNonfood
OrangeJuiceNotFrozen	HoProdFood	FloorWaxPolish	HoProdNonfood
PackagedFrozenRefrigeratedPasta	HoProdFood	FurniturePolish	HoProdNonfood
PackagedInstantPotatoes	HoProdFood	GlassAndSurfaceCleaners	HoProdNonfood
PackagesOfMiniatureCandy	HoProdFood	Glue	HoProdNonfood
PeanutButter	HoProdFood	HouseholdCleaners	HoProdNonfood
PickleRelish	HoProdFood	InBowlToiletBowlCleaners	HoProdNonfood
Pickles	HoProdFood	InTankToiletBowlCleaners	HoProdNonfood
PizzaMixesAndSauces	HoProdFood	IndoorInsecticides	HoProdNonfood
PizzaShellsCrusts	HoProdFood	IndoorPlantFood	HoProdNonfood
PoppingCornPopcornSnacks	HoProdFood	InsectRepellents	HoProdNonfood
Pretzels	HoProdFood	LaundryPreTreatmentsPreCleaners	HoProdNonfood
PuddingsPieFillings	HoProdFood	LightBulbs	HoProdNonfood
ReadyToEatDoughnuts	HoProdFood	Luggage	HoProdNonfood
ReadyToEatMuffins	HoProdFood	OutdoorInsecticides	HoProdNonfood
ReadyToEatSweetRollsPastries	HoProdFood	OvenCleaners	HoProdNonfood
RefrigeratedFrozenBreadAndDoughProducts	HoProdFood	PaintStain	HoProdNonfood
Rice	HoProdFood	PaperNapkins	HoProdNonfood
RiceCakes	HoProdFood	PaperTowels	HoProdNonfood
SaladOrCookingOil	HoProdFood	PlasticGarbageBagsTrashCanLiners	HoProdNonfood
SaladToppings	HoProdFood	PlasticSandwichFoodStorageFreezerBags	HoProdNonfood
Salt	HoProdFood	RubberGloves	HoProdNonfood
SaltAlternatives	HoProdFood	ScouringCleansers	HoProdNonfood
Sausage	HoProdFood	SoapDetergentsForFineFabrics	HoProdNonfood
SeasoningsSpices	HoProdFood	SoapDetergentsForRegularLaundry	HoProdNonfood
Shortening	HoProdFood	ToiletPaper	HoProdNonfood
SnackCakes	HoProdFood	TransparentTape	HoProdNonfood
SourCream	HoProdFood	WritingInstrumentsBrands	HoProdNonfood
SoySauce	HoProdFood	CannedWetCatFood	HoProdPets
SpaghettiPastaSauce	HoProdFood	CannedWetDogFood	HoProdPets
SprayNonStickCookingProducts	HoProdFood	CatLitter	HoProdPets
SpreadCheeseCheeseSauce	HoProdFood	DogBiscuitsOrTreats	HoProdPets
StuffingMixesAndStuffingProducts	HoProdFood	FleaTickCareProductsForDogsCats	HoProdPets
Sugar	HoProdFood	PackagedDryCatFood	HoProdPets
TableSyrupMolasses	HoProdFood	PackagedDryDogFood	HoProdPets
ToasterProducts	HoProdFood	HotelsMotelsWhereStayed	Hotels
Vinegar	HoProdFood	AutoInsurance	Insurance
		HomeownersOrPersonalPropertyInsuranceCompany	Insurance
		LifeInsuranceCompanies	Insurance
		MedicalInsuranceCompanies	Insurance
		FamilyRestaurantsSteakHouses	Restaurants
		FastFoodDriveInRestaurants	Restaurants
		ApplianceHardwareElectronicsStoresTimesShopped	Retail
		ConvenienceStoresTimesShopped	Retail
		DepartmentClothingSpecialtyStoresTimesShopped	Retail
		FurnitureStoresTimesShopped	Retail

Notes. MRI-Simmons product market names and sector names in our data.

Figure A.6: HHI over time, by sector and aggregation level

Notes. HHI over time, at the state group – product market (top), product market (second row), state group – sector (third row), and sector (bottom) levels. Each line shows the expenditure-weighted average of HHIs, for all local markets (top), markets (second row), and local sectors (third row) in a given sector. The left column shows results for manufacturing, the center column for food, beverage, and health products, and the right column shows results for non-manufacturing.