

DISCUSSION PAPER SERIES

IZA DP No. 10882

**Consumption Inequality and the Frequency
of Purchases**

Olivier Coibion
Yuriy Gorodnichenko
Dmitri Koustas

JULY 2017

DISCUSSION PAPER SERIES

IZA DP No. 10882

Consumption Inequality and the Frequency of Purchases

Olivier Coibion

University of Texas at Austin and NBER

Yuriy Gorodnichenko

University of California, Berkeley, NBER and IZA

Dmitri Koustas

University of California, Berkeley

JULY 2017

Any opinions expressed in this paper are those of the author(s) and not those of IZA. Research published in this series may include views on policy, but IZA takes no institutional policy positions. The IZA research network is committed to the IZA Guiding Principles of Research Integrity.

The IZA Institute of Labor Economics is an independent economic research institute that conducts research in labor economics and offers evidence-based policy advice on labor market issues. Supported by the Deutsche Post Foundation, IZA runs the world's largest network of economists, whose research aims to provide answers to the global labor market challenges of our time. Our key objective is to build bridges between academic research, policymakers and society.

IZA Discussion Papers often represent preliminary work and are circulated to encourage discussion. Citation of such a paper should account for its provisional character. A revised version may be available directly from the author.

ABSTRACT

Consumption Inequality and the Frequency of Purchases*

We document a decline in the frequency of shopping trips in the U.S. since 1980 and consider its implications for the measurement of consumption inequality. A decline in shopping frequency as households stock up on storable goods (i.e. inventory behavior) will lead to a rise in expenditure inequality when the latter is measured at high frequency, even when underlying consumption inequality is unchanged. We find that most of the recently documented rise in expenditure inequality in the U.S. since the 1980s can be accounted for by this phenomenon. Using detailed micro data on spending which we link to data on club/warehouse store openings, we directly attribute much of the reduced frequency of shopping trips to the rise in club/warehouse stores.

JEL Classification: D31, E21, D63

Keywords: consumption inequality, expenditure inequality

Corresponding author:

Yuriy Gorodnichenko
530 Evans Hall #3880
Department of Economics
University of California, Berkeley
Berkeley, CA 94720-3880
USA
E-mail: ygorodni@econ.berkeley.edu

* We are grateful to Per Krusell, Arthur Middlebrooks at the Kilts Center for Marketing, Taylor Wilson at the Bureau of Labor Statistics, as well as seminar participants at UT Austin, Houston, Alberta, and Montreal for helpful comments. We thank Marc Dordal-i-Carreras and Chaewon Baek for excellent research assistance. Gorodnichenko thanks the NSF for financial support.

Introduction

Income inequality has been rising sharply since the 1980s, raising concern among economists, policymakers, and the general public. However, whether consumption inequality has gone up in similar fashion, which is arguably more relevant for welfare, remains the subject of heated debate in the literature. Understanding what has happened to consumption inequality can also be informative about the forces underlying the rise in income inequality.

For example, Krueger and Perri (2005) argue that improved financial intermediation has allowed households to more easily smooth their consumption over transitory income shocks, thereby compressing consumption inequality. Relatedly, Pistaferri, Blundell and Preston (2008) argue that much of the rise in income inequality since the mid-1980s came from transitory shocks (as opposed to permanent shocks) that households are able to partially insure themselves against, consistent with the Permanent Income Hypothesis (PIH), and the absence of a commensurate rise in consumption inequality. On the other hand, Battistin (2003), Attanasio et al. (2007), Attanasio et al. (2015), Aguiar and Bils (2015) and others argue that the flat profile of consumption inequality is nothing more than a measurement artifact and that consumption inequality has risen in line with income inequality since 1980.

We build on this literature by emphasizing the distinction between *spending* (*expenditure*) inequality and *consumption* inequality. While households can enjoy a smooth consumption flow from most goods, their purchases may occur only infrequently. Because household surveys typically track expenditures for a short period of time to minimize recall error and reporting burden (e.g., a two-week period in the Diary Survey of the Survey of Consumer Expenditures (CEX)), measures of spending inequality can fail to correctly measure the underlying consumption inequality due to the timing of purchases. This can matter not only in the cross-section (if one household happened to buy paper towels in a period and another did not, spending inequality over that period would be higher than consumption inequality even if both households have the same flow consumption of paper towels) but also for measuring trends over time. To see the latter, suppose that consumers start stocking up on food once a month rather than once a week. Even if they maintain the same consumption flow, the cross-sectional inequality of spending measured at a less than monthly horizon will rise despite the fact that underlying *consumption* inequality would have stayed the same. In this paper, we document such a decline in the frequency of shopping, quantify its potential implications for historical changes in consumption inequality, and study its potential sources.

Our starting point is the well-documented difference in the trend of expenditure inequality across the Diary and Interview surveys of the CEX. While the latter points toward little change in expenditure inequality over time (as documented in Krueger and Perri 2005), the former instead suggests that expenditure inequality has risen more closely in line with income inequality (see e.g. Battistin 2003). While there are many potential sources for this difference, one is the differing frequency over which expenditures are measured: bi-weekly in the Diary Survey and monthly (or quarterly for some categories) in the Interview Survey.¹ Consistent with the frequency of expenditure measurement being a force behind the different inequality trends in the two surveys, we then document that the frequency of shopping has indeed systematically declined over time. Using data from the CEX Diary Survey, we find that the fraction of days in which households engage in any shopping for non-durable goods has been falling over time, so that households concentrate their shopping into fewer days of the week. Using even more detailed information on household expenditures from the Nielsen HomeScan data, we again document a decline in the number of days in which households do their shopping.² Hence, part of the greater increase in inequality as measured by the Diary Survey may indeed be coming from a changing frequency of shopping by households.

We provide several additional pieces of evidence, based on micro data, consistent with this conjecture. First, while average real expenditures on goods in the Nielsen sample has been approximately constant between 2004 and 2014, this masks underlying changes along the intensive and extensive margins of shopping behavior. The number of shopping trips (extensive margin) has been steadily falling over the entire sample, whereas the average expenditures per trip (intensive margin) have been rising. Hence, we see households making fewer, but larger, shopping trips on average. Second, using information on the volumes and sizes of individual goods purchased in the Nielsen sample, we find that households have been purchasing larger quantities or volumes of goods over time, consistent with increased stocking up. Third, using the American Time Use Survey, we compute average shopping times for individuals. We find a strong decline in the average amount of time spent shopping by US households, driven entirely by the extensive margin. Households do fewer trips per day and are less likely to go to any store on any given day. In

¹ Attanasio et al. (2007) provide a discussion of other potential sources for differing trends in expenditure inequality across the two CEX surveys.

² We first perform a battery of checks to ensure that the Nielsen data are comparable to CEX Diary Survey. We find that mean expenditures and implied inequality levels are quite close across the two datasets, once one focuses on goods that are common across the two.

contrast, the average duration of a shopping trip (the intensive margin) has held steady over this time period. These are precisely the expected patterns as households buy larger quantities of goods while at the store and therefore need to go to the store less often. Thus, the ability to stock up appears to be a critical component of these differences in trends.

To quantify the contribution of changing frequency of shopping to the differential trends in expenditure inequality across frequencies of aggregation, we pursue two approaches. Ideally, one would simply vary the duration of periods over which each household's expenditures are aggregated then construct cross-sectional measures of dispersion for each frequency. Unfortunately, the data in each of the CEX surveys are inadequate for this since households in the Diary Survey only report their expenditures for two weeks while households in the Interview Survey report their expenditures over one-to-three month periods but do not provide higher frequency variation within those periods. However, the Nielsen data tracks spending by households daily for extended periods, thereby allowing us to assess the extent to which trends in expenditure inequality are sensitive to the frequency over which expenditures are aggregated, e.g. weekly, biweekly, monthly, quarterly and annual. This approach yields five different measures of inequality based on differing time frequency aggregations, albeit over a more limited sample than the CEX. We use these series to assess the extent to which time aggregation affects the trend in spending inequality and document a clear effect of time aggregation on the trends in spending inequality. Short time horizons for measuring consumption yield positive trends in inequality but much flatter profiles at lower frequencies.³ When household spending is aggregated over the course of the year, there is essentially no trend in inequality. Hence, time aggregation can effectively account for all of the difference in the trends of consumption inequality identified by these two surveys.

Our second and complementary approach to quantifying these effects exploits the changing dispersion in individual households' expenditures over time (i.e., dispersion of a household's expenditures over the course of a year when expenditures are measured weekly, biweekly, etc.). We show using a simple model that the latter measure provides a convenient metric to assess the contribution of shopping frequencies to the cross-sectional dispersion of expenditures. Based on how individuals' time-dispersion of expenditures have changed over time according to the Nielsen

³ We perform a similar test in the CEX Diary survey by comparing trends in inequality of expenditures summed at the weekly vs biweekly frequency and in the CEX Interview survey by comparing trends in inequality of expenditures summed at the quarterly vs annual frequency. In each case, we find the same qualitative result that higher frequencies of aggregation lead to steeper trends in expenditure inequality.

data, we can attribute all of the differential increase in cross-sectional expenditure inequality across the two surveys to a changing frequency of shopping.

There are many mechanisms which could explain why American households are increasingly purchasing larger quantities when shopping and therefore shopping less frequently. One such mechanism is the rise of club/warehouse stores (Costco, Sam's Club, BJ's, etc.) which, by design, sell larger quantities of goods to households at lower unit prices. As these stores have expanded throughout the country since the 1980s, it has become easier for households to stock up in ways that were not feasible in the past, consistent with the decreased frequency of shopping that we observe. Furthermore, there is considerable geographic variation in the ease with which households can access one of these retailers, enabling us to quantify the contribution of this mechanism.

To assess whether club/warehouse stores can explain some of the rising concentration in household shopping trips, we characterize the link between how much variation there is in an individual's spending over time and their reliance on club stores in their expenditures. Specifically, we first measure the variation in a household's expenditures over a year using different time frequencies for subperiods: weekly, bi-weekly, monthly and quarterly. Households who do more infrequent shopping trips have relatively higher dispersion in their expenditures at higher frequencies than at lower frequencies. This greater time-series dispersion in expenditures for one household when it does large purchases infrequently is therefore analogous to how cross-sectional dispersion in consumption is higher when more households engage in infrequent shopping. To assess how much households use club stores, we measure the fraction of a household's expenditures that were spent at a club store over the course of that year.

The link between shopping at club stores and stocking up can then be assessed by regressing an individual's expenditure dispersion on that individual's share of expenditures going to club stores, using dispersion measures at different time frequencies. The results suggest that shopping at club stores is indeed correlated with significantly more stocking up. There is a strong positive correlation between the coefficient of variation at the weekly frequency and a household's share of expenditures at club stores, but this correlation declines rapidly as we increase the amount of time over which expenditures are aggregated, as expected since it becomes progressively more difficult to stock up for longer periods. Shopping at club stores also explains a diminishing fraction of the variance in households' coefficients of variation at longer durations for time aggregation of expenditures. An instrumental variable strategy, based on the differing distance of households from club/warehouse

stores, supports causality running from access to club/warehouse stores to increased stocking up in expenditures. We find that the increased prevalence of club/warehouse stores since the early 1980s can account for approximately 40 percent of the rise in measured inequality in expenditures.

This paper relates to a growing literature on measuring and understanding the sources of economic inequality. Unlike much recent work on the rising share of income and wealth of the top 1% (e.g., Piketty and Saez 2003, Piketty et al. 2016), we focus on inequality outside of the top 1% since our data sources are not informative about top income earners. Instead, our results build on the literature relating consumption and income inequality amongst households in the bottom 99% in the U.S. (e.g. Krueger and Perri 2005, Aguiar and Bils 2015, Attanasio and Pistaferri 2016) or abroad (Gorodnichenko et al. 2009). Relative to these papers, our contribution is to document how the changing frequency of shopping contributes to the *measured* trends in spending inequality. This paper is also closely related to a growing literature on shopping behavior of households and its implications for macroeconomics. For example, Nevo and Wong (2015) focus on the substitution between expenditures and home production during the Great Recession. Wong (2016) studies infrequent purchases of durable goods, Aguiar and Hurst (2013) focus on life-cycle consumption patterns of households, and Coibion, Gorodnichenko and Hong (2015) study the implications of store-switching for macroeconomic dynamics. We extend this line of work to study how shopping behavior can affect the measurement of expenditure inequality. Another closely related line of work focuses on household search for goods. Menzio and Trachter (2015) for example focus on the implications of consumer search for equilibrium price dispersion while Michailat and Saez (2015) study the implications of product market search for macroeconomic dynamics. Finally, our use of time use data to study household decisions echoes Aguiar et al. (2013) and Lee et al. (2012). But whereas they use time use surveys to study the substitution between labor, leisure and home production, we exploit information on time spent shopping and traveling to stores to characterize the changing nature of household shopping behavior in the U.S.

The paper is organized as follows. In section 1, we discuss the different CEX surveys and their implications for trends in expenditure inequality, as well as providing preliminary evidence on how the frequency of shopping has declined over time. Section 2 introduces the Nielsen data and provides additional evidence on the changing characteristics of household shopping. Section 3 provides two ways of quantifying the contribution of these changing shopping patterns to

expenditure inequality trends. Section 4 assesses how much of the changes in household shopping behavior can be attributed to the growing prevalence of club stores. Section 5 concludes.

I. Expenditure Inequality and the Changing Frequency of Purchases

To measure consumption inequality, previous work such as Krueger and Perri (2005), has focused primarily on the Consumer Expenditure Survey (CEX). Since the CEX is a well-known and well-documented data source, we provide only a brief overview of these data. We focus in particular on the differences between the two main components of the CEX, both of which have been used to measure consumption inequality: the Interview Survey (IS) and the Diary Survey (DS). We also highlight changes in the survey methodology over time that could impact the dispersion of measured spending.

In the IS, about 1,500-2,000 households are asked each month to *recall* the dollar value of spending over the previous month or quarter (depending on the category). Households are interviewed once per three months for five consecutive three-month periods, although the BLS only makes data available for interviews two through five. While early Interview surveys exist in 1960-1 and 1972-3, the modern Interview Survey begins in 1980 and is not directly comparable to prior waves of the IS. In the 1981-1983 waves for the IS and 1982-3 waves for the DS, only urban households were sampled due to budget cuts. A main advantage of the Interview Survey is its broad coverage of goods purchased by households (approximately 95% of typical household's consumption expenditures) since it is used to create expenditure weights for the Consumer Price Index.

A separate sample of households participate in the Diary Survey. Households are asked to record their spending each day for two weeks in a diary, which is later transcribed by U.S. Census Bureau officials. Records of daily spending become available to researchers starting in 1982, for the categories of food-at-home as well as food away from home. In 1986, the Diary Survey was expanded to cover a comprehensive set of spending categories. In 2004, the Census Bureau adopted a variety of changes to Diary data collection procedures that resulted in potentially more accurate recording of purchases, including computer assisted technology for U.S. Census Bureau enumerators. In the figures made using the DS, we include a vertical line to indicate these structural breaks.

An extensive literature exists discussing the pros and cons of the two surveys. For example, Krueger et al. (2010), Aguiar and Bils (2015) and Attanasio et al. (2012) find that the Interview survey in the CEX underreports spending relative to aggregate data and that this underreporting

has become more severe over time. On the other hand, Bee et al. (2012) compare reported consumption spending data in the CEX to comparable data from the national income accounts data and find that the CEX data conform closely to aggregate data for large consumption categories. Battistin (2013) and Attanasio et al. (2007) argue that, given data in the DS, the IS underestimates the rise in expenditure inequality since the 1980s. In contrast to the view promoted by Krueger and Perri (2005) that expenditure inequality (measured using the CEX IS) has not risen nearly as much as income inequality, more recent work has instead concluded that expenditure inequality has in fact grown more rapidly than implied by the CEX IS.

To illustrate how pronounced the differences are between the Interview and Diary surveys are for resulting trends in expenditure inequality, we construct a coefficient of variation for each survey. Specifically, for each survey, we measure each household's expenditures on non-durable goods and services.⁴ In the Diary survey, expenditures are daily while in the Interview Survey they are over a monthly or quarterly horizon. We use BLS's monthly Personal Consumption Expenditures (PCE) price index to deflate household spending (for quarterly spending, we follow the BLS and allocate spending equally across months). Because the CEX IS and DS surveys use a somewhat different classification of goods and services (universal category codes, UCCs), we update the concordance created in Bee et al. (2012) and increase its scope to be comprehensive of all UCC codes (see Appendix E). We then calculate the coefficient of variation in expenditures across all households (the ratio of the cross-sectional standard deviation to the cross-sectional mean of expenditures) for each year. We use coefficient of variation (CV) as our baseline measure of inequality because CV allows us to include households reporting zero spending in a given period, which is important while working with high-frequency (e.g., weekly) data on spending. To minimize adverse effects of outliers on measures of inequality, we winsorize the right tail of household spending for a given frequency in each year at 1 percent.

The resulting time series are plotted in Figure 1. Using the Interview Survey, we replicate the baseline result of Krueger and Perri (2005), finding little increase in expenditure inequality between 1980 and 2015. In contrast, the Diary Survey reveals a pronounced increase in expenditure inequality from 1980 to the early 2000s. The ratio of the two inequality measures provides a simple way of examining differences in trends across the two: *this ratio is increasing systematically over*

⁴ Our coverage of non-durable goods and services follows Coibion et al. (2012). Clothing and most services are not consistently measured in the Diary survey until 1986.

time, going from 1.05 in 1980 to 1.35 in 2015. Its persistent increase even since the early 2000s reflects the fact that spending inequality as measured by the Interview Survey is declining over this time period but approximately constant in the Diary Survey.

This difference in implied inequality trends is not driven by composition effects, either in terms of composition of goods or characteristics of households. For the former, we can compare spending inequality in the two surveys for matched and consistently (over time) collected categories of goods, thereby controlling for potential changing compositions of purchases over time. We find that the same trend in the ratio of inequality across the two surveys holds (Appendix Figure A1). Similarly, we can control for potentially changing household characteristics by looking at residual inequality in each survey. We do so by regressing household expenditures on a large set of observable characteristics of households (age, income, etc.) in each survey, then construct equivalent inequality measures from the residuals of household expenditures:

$$Y_{ht} = \mathbf{X}_{ht}\boldsymbol{\gamma} + \epsilon_{ht} \quad (1)$$

where h and t index households (respondents) and years, Y is a variable of interest, \mathbf{X} is a vector of controls which includes a polynomial in the age of household head, gender dummy for household head, a set of race dummies for household head, a set of dummies for educational attainment of household head, number of children, dummy for employment of household head, and a set of region dummies. The coefficient of variation adjusted for changes in demographics is calculated as $\sqrt{\text{var}(\epsilon_{ht})}/\text{mean}(\hat{Y}_{h,t})$. The results yield a similar pattern of a systematically rising ratio of consumption inequality in the DS relative to the IS from 1.1 in 1980 to 1.5 in 2015 (see Figure 1). Results are also similar if we use different metrics for measuring inequality (Gini coefficients are plotted in Appendix Figure A3) or within subgroups of the population. For example, in Appendix Figure A4, we document that the same patterns of rising inequality in DS survey relative to IS survey hold within both younger and older households, for the employed and the non-employed, for different races, for households of different sizes, as well as for households of high and low income.

It is also worth noting the large difference in level between the two series. Although the IS likely has larger measurement error than the DS, this difference in levels is to be expected since the DS measures expenditures at the biweekly frequency whereas the IS measures expenditures over a monthly/quarterly horizon. Since some goods are purchased infrequently, the Diary Survey will record zero expenditures for some households and large expenditures for others depending on the timing of their purchases. In contrast, the Interview Survey will more consistently measure

positive expenditures due to the longer horizon. By the same logic, inequality among weekly household expenditures in the Diary Survey is approximately 20 percent higher on average than for expenditures at the bi-weekly frequency in the same survey (see Appendix Figure A2).

Importantly, the fact that expenditures are measured over different horizons can be a source of differences in *trends* of measured “consumption” inequality if the frequency of household purchases is changing over time. For example, if households change their frequency of purchasing toilet paper from a weekly to a monthly frequency while keeping their flow consumption of toilet paper unchanged, this would induce a rise in the ratio of spending inequality when expenditures are measured at the bi-weekly frequency relative to when expenditures are measured at the monthly frequency. In this case, consumption inequality would not have changed (everyone is still using the same amount of toilet paper per unit of time) but inequality in spending at high frequencies would rise thus underscoring the difference between spending inequality and consumption inequality.

There is evidence consistent with this hypothesis. Since the CEX diary survey provides daily expenditures, we can measure the average number of days in which households engage in positive expenditures (out of 14 possible days in the diary) for each survey year.⁵ The result is plotted in Figure 2. Panel A refers to all nondurables including services, while Panel B focuses on a more limited set of nondurable goods: food at home, alcohol, tobacco, and other small nondurable goods. In 1980, households purchased a positive amount of nondurable goods 9 out of the 14 days of each bi-weekly period, but this number had fallen to 8 days by 2004. There is a structural break in the series of 2005 (because of the changes in how CEX DS data are collected), but the average number of days falls another 0.5 by 2015. Panel B shows an even steeper drop in shopping for the narrower set of nondurables. Almost identical trends are obtained if we use positive values (e.g. \$5, \$10, etc. including with inflation adjustment) as the threshold for daily expenditures instead of zero.

Like the changing ratio of expenditure inequality in the two surveys, the declining frequency of shopping, at least as measured in days with positive spending, holds for a wide range of products and is not driven by household characteristics, such as a growing share of working spouses. To see the latter, we construct residual measures of the number of days with positive expenditures for each household after controlling for the household observable characteristics as before and measure the average across households (normalizing it to have the same value as the raw measure in 1980 and

⁵ Since the Interview survey does not provide high-frequency expenditure data, we cannot construct equivalent measures in that data.

again in 2005). The trends are almost identical, so the declining frequency of days with positive shopping experiences is not coming from changes in household characteristics.

Unfortunately, the CEX data present many limitations which do not allow us to characterize these effects in a more detailed way. For example, without more detailed information on households' shopping activities, we cannot quantify whether households are doing fewer shopping trips or are combining the same number of trips into fewer days. Without information on quantities and sizes of purchased goods, we cannot assess e.g. whether households are buying larger quantities on their less frequent trips. Without information on time use, we cannot determine whether households are changing the amount of time they devote to shopping. And because neither the DS nor IS has long panels of high-frequency data on expenditures, we cannot quantify the extent to which changing frequencies for computing expenditure inequality contributes to the differential trends in spending inequality. Using additional data sets, we explore these questions in the next section.

II. Changing Patterns of Household Shopping Behavior

Evidence from the CEX Interview Survey suggests that households have been reducing the frequency of their shopping over time, a feature potentially explaining the differing trends in spending inequality observed in the Interview and Diary Surveys. In this section, we provide more detailed microeconomic evidence on the changing patterns of household shopping behavior. To do so, we begin by introducing an additional but more detailed dataset on household shopping, the Nielsen Home Scanner data and show that these data display similar shopping characteristics as the CEX. For example, in addition to presenting comparable measures of expenditures on different categories of goods, we confirm the finding from the CEX that households are doing their shopping on fewer days using the Nielsen data. Second, we show that while average real expenditures by households were relatively constant between 2004 and 2014, households have been spending relatively more on each shopping trip but doing fewer shopping trips per year, consistent with increased stocking up. Third, using data on quantities, we show that households are purchasing ever larger sizes or numbers of units of individual products on each trip. Fourth, we turn to time use surveys to show that the amount of time households have been spending on shopping has been decreasing. All four facts are consistent with households increasingly “stocking up” in their shopping.

II.A Characteristics of the Nielsen data

While CEX allows us to construct time series going back to the early 1980s and have a good coverage of goods and services purchased by households, the data in each of the CEX surveys present some limitations. For example, households in the CEX Diary Survey only report their expenditures for two weeks, so we may be missing important expenditures that are not made over that two-week measurement period. Households in the CEX Interview Survey report their expenditures over one month (or three months depending on the category) but do not provide higher frequency variation within those periods. Because the Diary and Interview surveys are not connected in any way, we cannot establish how time aggregation affects trends in measured spending inequality. Moreover, measurement error due to by recall bias in the IS imposes additional challenges when comparing consumption inequality at quarterly frequency in the IS versus higher frequency in the DS.

To address these challenges, we turn to Nielsen Home Scanner (Nielsen) data, available through the Kilts Center at the Booth School of Business at the University of Chicago, which provides a source of rich, high-frequency household spending data. Nielsen data are currently available from 2004 to 2014. From 2004-2006, the sample included approximately 40,000 households, increasing to 60,000 households beginning in 2007. Over the period 2004-2014, the mean and median tenure in the sample were approximately 4 and 3 years, respectively.

The Nielsen sample is comprised of a combination of households recruited by Nielsen, as well as unsolicited volunteers. In exchange for their participation, households receive points that can be redeemed for prizes as well as entry into lotteries that award more points or cash. Households are provided a scanning device by Nielsen to scan the barcodes of their purchases and they are encouraged to scan newly purchased items as soon as they return home. Nielsen employs their own sample filter, requiring that households must report a minimum dollar amount per month, which varies depending on household size, to be in the final sample. To ensure that our results are not driven by households with incomplete records, we include only households with a least one shopping trip where they scan items in each month of a given year.

After scanning a product using the device, households directly report the quantity of the barcode that they purchased. For a group of participating stores, prices are automatically reported to Nielsen; otherwise the household is also asked to manually enter the product price. Nielsen later merges in information about the product that is tied to the barcode, including a measure of volume

or count if applicable. If a product does not have a barcode, a purchase of this product is generally not reported as the main Nielsen data focus primarily on nondurables with a barcode.

Household demographics, including zip code and employment status, are updated once per year as part of a household survey. Nielsen uses the demographic information to construct household weights that weigh the sample to be nationally representative. The household spending data are technically available on a daily basis. However, in some cases, the purchase date in Nielsen could reflect the date the data were transmitted by the scanning device to Nielsen, rather than the true purchase date by the household.⁶

The Nielsen data include over 325 million barcodes that Nielsen estimates to cover approximately 30 percent of household spending. These barcodes are categorized by Nielsen into lower levels of aggregation. Nielsen's "Product Groups," of which there are 125, are closest to universal classification codes (UCCs) in the CEX Diary Survey. For our analysis, we construct a correspondence table between CEX Diary UCCs and Nielsen Product Groups (see Appendix B).

To ensure that our results are not driven by the specifics of how Nielsen data are collected, we compare basic moments for categories of consumer spending in CEX Diary Survey and in Nielsen. All statistics are for the biweekly frequency. We compute moments for categories of goods present in both sources so that the coverage of goods is comparable across data sets (e.g., Nielsen data have virtually no coverage of services). The set of comparable non-durable goods generally includes food, alcohol, and small non-durables. To differentiate the frequency of shopping trips and the size of purchases, we show the share of households reporting zero spending over two weeks for a given category of goods ("zero share") and moments (mean, standard deviation, interquartile range) for the size of purchases conditional on a purchase in the category. Results for selected categories of goods for year 2014 are reported in Table 1.

Consistently across data sources, we observe that purchases for many categories of goods are not made frequently. On average, there is an approximately 80 percent chance that there is no purchase in a typical category of goods over two weeks. Furthermore, for the comparable categories, the probability of no purchases for any of the categories during the period is 6 percent

⁶ Nielsen made changes in 2009 that resulted in more purchases being assigned a transmission date rather than the true purchase date. We therefore must be cautious comparing higher frequency (especially daily) behavior across these regimes.

in the CEX data and 10 percent in the Nielsen data. The correlation of zero shares across the surveys is 0.74 thus indicating high consistency across data sources.

Average spending conditional on a shopping trip is higher in the CEX than Nielsen data. For example, the average total bi-weekly spending on comparable categories of nondurable goods is \$239 in the CEX data compared with \$149 in the Nielsen data. This difference reflects the fact that the Nielsen data report considerably lower levels of spending for categories with few UPCs such as “fresh meat”, “fresh produce”, “bread and baked goods”, and “lawn and garden.” The levels of spending are much closer for categories of goods populated by products with UPCs. For example, the average spending on “baby food” (a category where most goods have UPCs) in the Nielsen data is close to the average spending on “baby food” in the CEX data while the average spending on “fresh meat” in the Nielsen data is smaller than the average spending on “fresh meat” in the CEX data. Note, however, that despite this limitation, the correlation between average spending or dispersion of spending in the CEX and Nielsen data is above 0.85 and can be further increased (up to 0.95) if a few outlier categories such as “fresh meat” are excluded. Measures of dispersion across the sources are close to each other too. We conclude that Nielsen data provide a useful complement to the CEX Diary Survey data for an analysis focusing on nondurable goods.

II.B Evidence from the volume/size of purchases

We can assess the possibility of increased stocking up by looking directly at volumes purchased: we should not only see increased spending per trip, but also increased physical volumes of goods purchased by households. Because Nielsen data report not only dollar spending for each universal product code (UPC, a precise definition of a good) but also units purchased as well as volumes of units, we can check if this prediction is borne out by the data. In particular, the Nielsen dictionary of UPCs specifies count or weight for each UPC. Using 2004 as the benchmark year, we examine the distribution of purchased weights or counts for each module⁷ of goods. We identify a purchase as “large volume” if the purchased weight or count is greater than the 90th percentile of the size distribution of purchased weights or counts in the module. Then for each year we compute the share of purchases (by weight or count) for each module and household holding the “large volume”

⁷ A module in the Nielsen data is a highly-disaggregated category of goods. There are over 1,000 modules in the data. Example of modules are “FRUIT JUICE - APPLE”, “FRUIT JUICE - GRAPE”, “MEXICAN SHELLS”, “MEXICAN TORTILLAS”, “DAIRY-MILK-REFRIGERATED”, “DAIRY-BUTTERMILK-REFRIGERATED”, “DAIRY-CREAM-REFRIGERATED”, “EGGNOG - FRESH & CANNED”.

threshold constant across years. Using annual expenditure shares to aggregate across modules and sampling weights to aggregate across households, we construct an average share of “large volume” purchases for each year. We find (Figure) that over time “large volume” purchases were increasingly prevalent. For products sold by weight, the share of large-volume purchases increased from 16.8 percent in 2004 to 22.3 percent in 2014. For products sold by count, the share rose from 19.3 percent to 21.8 percent over the same period. As before, these results are robust to conditioning on observable household characteristics or considering other thresholds for large-volume purchases (see Appendix Figure A7). They therefore provide direct evidence that households are engaging in larger-sized purchases of goods.

II.C Evidence from the intensive and extensive margins of purchases

In addition to the evidence from inequality measures and days of with positive shopping, we can assess more directly whether households are indeed stocking up more on their increasingly infrequent shopping trips. One indicator of shopping becoming increasingly concentrated over time would be that the amount of spending per shopping trip should be increasing as long as the total annual spending is stable. To assess this prediction, we decompose average annual expenditures by households into their average expenditure per shopping trip (the intensive margin) and their average number of shopping trips per year (the extensive margin). The results are presented in **Error! Reference source not found.** The figure shows three lines: average log annual spending per household, the average number of shopping trips per year, and the average log spending per shopping trip. All series are normalized to be equal to one in 2004. While annual spending is approximately constant over 2004-2014, we see that the number of shopping trips declines by close to 20 percent while the average spending per trip increases by the same amount. Hence, households are doing fewer shopping trips, which is consistent with the CEX data, but spending more on each trip. Furthermore, the results are unchanged if we control for household observables as in section I, so these patterns are not driven by a changing composition of households but rather by changing behavior of households.

II.D Evidence from time spent shopping

Our argument suggests that households should increasingly buy goods in bulk and consequently spend less time shopping. While Nielsen data do not permit us to assess changes in shopping time

for purchases of goods (e.g., we know the number of shopping trips but not their duration), we can use the American Time Use Survey (ATUS) to examine the evolution of households' shopping time.

Since 2003, U.S. Census Bureau on behalf of the Bureau of Labor Statistics (BLS) surveys a randomly chosen subset of households participating in the Current Population Survey (CPS) to report their time use for a given day. Each year, approximately 25,000 households are requested to recall their activities for a 24-hour period and provide detailed information on the type and duration of each activity.⁸ For each activity start/end times are indicated which allows us to observe how many shopping trips were done by a respondent. Time spent for purchases of goods includes not only shopping time but also travel time, researching time, comparison time, etc. Because ATUS respondents are sampled from the CPS, we also have detailed demographic information (age, gender, marital status, educational attainment, employment status, income bracket, etc.).

Using this information, we compute average shopping time for each year and report the resulting series in Figure . In addition to the total time spent on purchases of goods, we show the intensive (average time per shopping trip) and extensive (number of shopping trips per day) margins of shopping. Because the composition of U.S. population has been changing over time, we also present series adjusted for the changes using a specification similar to regression (1). The type of regression (1) depends on the nature of the dependent variable. When the dependent variable is the number of trips, we use a Poisson regression. For the average time spent on shopping for purchases of goods we use a Tobit regression (because the distribution is censored at zero). For the average duration of shopping trips (which is conditional on having a trip), we use OLS. When the dependent variable is an indicator variable for having a shopping trip on a given day, we use a logit regression. In cases other than OLS, we take $\hat{\beta}$ as the marginal effects calculated at means.⁹

Figure 5 documents that shopping time (Panel A) has been declining since 2003. Adjusting for the observed characteristics of respondents yields an even greater decrease. Panels B through D show that this reduction in shopping time is driven exclusively by the extensive margin rather

⁸ There are precursors of the ATUS. An early time-use survey was implemented in 1965. Subsequent time-use surveys were done in 1975, in the mid-1980s and in the mid-1990s. Unfortunately, these earlier surveys differ in sample design, coverage and level of detail. To ensure consistency of the series, we restrict our analysis to the surveys implemented by the BLS since 2003.

⁹ In addition to demographic characteristics of households, these regressions include a set of dummies to capture within-week variation in shopping intensity and income brackets. We include dummies for week days because the day-of-week sampling of respondents has changed over time. We include controls for income because as income increase households may engage in more shopping (shopping is leisure) or less shopping (shopping is home production). We generally find that in the cross-section high incomes are associated with higher shopping time.

than the intensive margin. Indeed, the average duration of a shopping trip (Panel D) varies over time but does not exhibit any trend. In contrast, the probability of having a shopping trip (Panel C) and the number of trips (Panel B) decline over time. These patterns are consistent with households doing fewer shopping trips but increasing the sizes of the products they buy during these trips. We find similar patterns for demographic subsamples (see Appendix Figure A5). In addition, the decline in shopping trips with no rise in shopping time (Panel D) suggests that households are engaging in larger purchases at approximately the same number of stores, not combining multiple store visits into single trips. As a result, it is unlikely that increasing geographical concentration of stores into strip malls, shopping centers, etc. that lowers the fixed costs associated with a shopping trip can account for increased lumpiness of consumer purchases.

III. Quantifying the Contribution of Changing Shopping Patterns

Given this evidence on how shopping patterns have changed over time, we are interested in quantifying the potential contribution of this channel to explaining the differential trends that we observe across survey measures of expenditure inequality. We consider two ways of doing so. The first exploits the fact that, in the Nielsen data, we track high-frequency expenditures of households over extended periods of time and can therefore construct measures of expenditure dispersion that mimic the timing of the CEX Interview and Diary surveys, as well as even shorter and longer periods of aggregation. This provides a direct test of how the frequency of aggregation can affect measured trends in expenditure inequality. The second method relies on the fact that we can also measure the dispersion in an individual’s expenditures over time, which will be directly related to their frequency of shopping. We show using a simple model that this “time-dispersion” in expenditures can contribute to the measured cross-sectional inequality in expenditures and that it can be used to quantify the contribution of changes in shopping patterns to trends in cross-sectional expenditure inequality.

III.A Trends in expenditure inequality

With the Nielsen data, we can examine directly how spending inequality varies with the level of time aggregation. Let X_{htpl} be spending of household h in period p (a week, bi-week, month, quarter of a given year, or a year itself) of calendar year t in location l (zip code, metropolitan area, or national level). Suppose the frequency of p is set to a week. Then for each week p of year t , we calculate the cross-sectional coefficient of variation $CV_{tl} = \sigma_{tl}/\bar{X}_{tl}$ where average spending

for period p year t is $\bar{X}_{tl} = \frac{1}{\#h} \sum_h \frac{1}{\#p} \sum_h x_{htpl}$ and the standard deviation of spending for the period is $\sigma_{tl}^2 = \left(\frac{1}{\#h \times p} \sum_{hp} (x_{htpl} - \bar{X}_{tl})^2 \right)$. The procedure for other frequencies is similar. In the Nielsen data, we treat weeks with no shopping activity over p as a true “zero”; in the DS and IS, we only keep households that have complete records over the length of p (two weeks of diaries for analysis at biweekly frequency, and four quarters of interviews for analysis at the annual frequency). Our measure of spending in the Nielsen data includes three major categories of goods: food-at-home, alcohol/tobacco, and small non-durables (e.g., paper towels, razors).

Panel A of Figure 6 plots the resulting measures of consumption inequality using the different levels of time aggregation for the 2004-2014 period. We observe two important patterns in the data. First, as we increase the level of time aggregation, the level of spending inequality declines. For example, the coefficient of variation for the weekly frequency is between 1 and 1.2 while for the biweekly frequency it is approximately 0.8. At the annual frequency, the coefficient of variation is less than 0.6. If household consumption were equal to household spending, we should not have observed such dramatic differences. The inequality of spending decreasing in the level of aggregation is consistent with consumption being smoother than spending.

Second, the trends in expenditure inequality are different across frequencies. While spending inequality measured at high frequencies (weekly and biweekly) increases over time, it is generally flat when measured at low frequencies (quarterly and annual). Table 2 reports the average annual change in inequality by frequency and documents that the slope of the time trend decreases considerably in the frequency of time aggregation until we reach the quarterly frequency of aggregation. Thus, simply changing the time horizon over which one measures expenditures significantly alters the measured growth in expenditure inequality, and in precisely the direction that we would expect if households are reducing the frequency at which they purchase goods. This difference in time aggregation could potentially account for much of the difference in observed trends between the Interview and Diary survey measures of expenditure inequality.

We can use disaggregated data to further explore this insight. Specifically, for each module in the Nielsen data for year 2014, we compute the ratio of spending inequality at the weekly frequency to spending inequality at the annual frequency. Then we relate this ratio to the frequency of shopping trips households have on average for goods in the corresponding modules.¹⁰ We find

¹⁰ This frequency is taken from Baker and Kueng (2017). We are grateful to Lorenz Kueng for sharing the data.

(Figure 7) a strong negative relationship between the ratio and the frequency of shopping, which is consistent with the predictions of our theory. Given that the frequency of shopping trips has declined, we can in principle reconcile why the levels and trends are different for spending inequality measured at different frequencies.

We can also directly explore the importance of the frequency of time aggregation for expenditures in the CEX data, albeit in a more limited way than in the Nielsen data. Within the Diary Survey, we can determine whether there is a difference in the growth of expenditure inequality when expenditures are measured bi-weekly, as done in Figure 1, versus an even higher frequency: weekly. Columns 1 and 2 of Panel B in Table 2 report the results: the growth in expenditure inequality in the Diary Survey is significantly larger when expenditures are measured at the weekly frequency than the bi-weekly frequency. Within the Interview Survey, we can compare trend growth in expenditure inequality measuring expenditures at the quarterly (three-month period) frequency versus the annual frequency. Columns 4 and 5 of Panel B in Table 2 report the results. As with the Nielsen data, we find no significant difference in the slopes, suggesting that few purchases in these data are conducted at a less than quarterly frequency. We reach the same conclusions when we restrict the CEX data to include goods that are covered in our Nielsen sample (Panel C). In short, the Nielsen data provide additional evidence that most of the difference in expenditure inequality trends observed between the Interview and Diary surveys can be accounted for by time aggregation of expenditures.

III.B The dispersion over time of a household's expenditures

A related approach to quantifying the contribution of time aggregation of expenditures to cross-sectional inequality is to consider the time-dispersion of households' expenditures. When households make their purchases less frequently, we will observe rising dispersion in a household's expenditures when those expenditures are measured over sufficiently short periods. Hence, we should observe similar patterns in the time dispersion of expenditures across frequencies of aggregation as we do in the cross-sectional data.

To see more precisely how the time-series dispersion of expenditures for individuals relates to the cross-sectional dispersion of expenditures, consider an environment similar in spirit to the celebrated Baumol-Tobin model. Specifically, each household h consumes a target dollar amount of consumption C_h over a total period of time of length T (e.g. a year where $T = 52$ weeks). Suppose

a household makes equally-sized purchases only on N_h periods out of the T . In a period when the household makes a purchase (which happens N_h/T of the time), that purchase $X_{h,t}$ is $X_{h,t} = C_h/N_h$ while in other periods (the remaining $1 - N_h/T$ fraction of the time) its purchases are $X_{h,t} = 0$.

We assume households can smooth their consumption over time so that, regardless of N_h , their per-period consumption is $\bar{C}_h \equiv C_h/T$. In this case, the cross-sectional average level of consumption across all T periods is $E_h[\bar{C}_h] \equiv \bar{C}$ and the corresponding dispersion of consumption, as measured by the cross-sectional coefficient of variation is $CV_h(\bar{C}_h) \equiv \sqrt{\text{var}(\bar{C}_h)}/\bar{C}$, which we take as given. Note that, if one measures dispersion in *consumption* for a given household over each of the T sub-periods, it will be identical to the dispersion in consumption over a longer time period because by assumption households can perfectly smooth their consumption flow. If we could measure consumption flow directly, the time horizon used for measuring those flows would not matter for the resulting measures of the cross-sectional dispersion of consumption.

To see the link between the cross-sectional dispersion in expenditures and the time-variation in each household's expenditures, it's helpful to start with the latter. Suppose we measure expenditures for each of the T subperiods for household h . The average expenditures across T subperiods for household h is $E_T(X_{h,t}) = \frac{C_h}{N_h} * \frac{N_h}{T} + 0 * \left(1 - \frac{N_h}{T}\right) = \frac{C_h}{T} \equiv \bar{X}_h$ and the variance of these expenditures for household h over the T periods is $\text{var}_T(X_{ht}) = \frac{N_h}{T} \left(\frac{C_h}{N_h} - \bar{X}_h\right)^2 + \left(1 - \frac{N_h}{T}\right) (0 - \bar{X}_h)^2 = \bar{X}_h^2 (T/N_h - 1)$. Hence, the coefficient of variation for household h when its expenditures are measured over subperiods is given by $CV_T(X_{h,t}) \equiv \sqrt{\text{var}_T(X_{ht})}/\bar{X}_h = \sqrt{T/N_h - 1}$ so that a household's time dispersion in expenditures is directly related to its frequency of shopping. As a household increasingly bunches its expenditures into fewer shopping trips (N_h falls), the time-series dispersion in its measured expenditures will rise.

To measure time-series variation in purchases in the Nielsen data, we follow our previous notation and let X_{htpl} be spending of household h in period p (a week, bi-week, month, quarter of a given year) of calendar year t in location l . We calculate the average per period spending for household h in year t as $\bar{X}_{htl} = \frac{1}{\#p} \sum_{pet} X_{htpl}$ and the standard deviation of spending for household h in year t as $\sigma_{htl}^2 = \left(\frac{1}{\#p} \sum_{pet} (X_{htpl} - \bar{X}_{htl})^2\right)$ and compute each household's coefficient of variation for spending over the course of the year as $CV_{htl} = \sigma_{htl}/\bar{X}_{htl}$. Households who do more infrequent

shopping trips have relatively higher standard deviation in their spending at higher frequencies than at lower frequencies. This greater time-series dispersion in spending for one household when they do large purchases infrequently is therefore analogous to how cross-sectional dispersion in consumption is higher the more households engage in infrequent shopping. We calculate CV at four frequencies: weekly, bi-weekly, monthly and quarterly.

In Panel B of Figure 6, we plot time series of the average (across households) coefficient of variation of each household's expenditures over time (that is, $\overline{CV}_{tl} = \frac{1}{\#h} \sum_h CV_{htl}$), using different time frequencies ranging from weekly ($T=51$ weeks)¹¹ to quarterly ($T=4$ quarters). As expected from infrequent purchases, the dispersion in households' expenditures is higher on average at high frequencies of aggregation and as the frequency of aggregation declines, the dispersion in expenditures falls toward zero. For example, the average time-series CV at weekly frequency is about 1, while the average time-series CV at the monthly frequency is approximately 0.4. Importantly, we can see a rising trend in the time dispersion of households' expenditures, particularly at higher frequencies, which is consistent with what we would expect when N is falling over time. The results are similar when we control for household characteristics (see Appendix C).

To relate the cross-sectional and time-series measures of dispersion, note first that the coefficient of variation in *expenditures* summed across all T periods is equal to the cross-sectional dispersion in *consumption*: $CV_h(\bar{X}_h) = CV_h(\bar{C}_h)$. That is, by using a long period over which to aggregate expenditures, one can recover the underlying dispersion in consumption. Now suppose instead we measure expenditures each subperiod. Then one can show (see Appendix C for derivations) that the cross-sectional coefficient of variation for expenditures across households at a given frequency (weekly, biweekly, etc.) is

$$CV_h(X_{h,t}) \approx CV_h(\bar{X}_h) \sqrt{1 + \left(\frac{1}{CV_h(\bar{X}_h)^2} + 1 \right) \{ \overline{CV_T(X_{h,t})} \}^2}$$

where $\overline{CV_T(X_{h,t})} \equiv E_T CV_h(X_{h,t})$ is the average across households of the time-dispersion of expenditures for each household at the given frequency. The first term ($CV_h(\bar{X}_h)$) captures the fact that higher underlying cross-sectional dispersion in consumption will lead to a higher dispersion of measured expenditures, even when the latter are measured at a higher frequency. For this term,

¹¹ In the Nielsen sample design, households exiting the sample do not have observations in the last few days of the calendar year. To ensure this does not affect our results, we focus on the first 51 complete weeks of the year (or 50 weeks for biweekly frequency).

we use $CV_h(\bar{X}_h)$ with \bar{X}_h measured over a long horizon (e.g. a year) to measure $CV_h(\bar{C}_h)$. The last term ($\overline{CV_T(X_{h,t})}$) reflects the fact that more dispersion in each individual's expenditures *over time* will lead to a higher level of dispersion in the *cross-section* as well, as long as $N_h < T$ for some h , i.e. as long as households' purchases are made less frequently on average than the length of the measurement period. As can be seen in Figure 6, increasing the duration over which expenditures are aggregated pushes the time dispersion of expenditures toward zero, but the cross-sectional dispersion is converging to a positive value, which according to the model is equal to the underlying dispersion of consumption.

Using average values of $\overline{CV_T(X_{h,t})}$ at the weekly frequency and $CV_h(\bar{X}_h)$ at the annual frequency in the Nielsen data yields $\partial CV(X_{h,t}) / \partial \overline{CV_T(X_{h,t})} \approx 1$.¹² Given that the time-series dispersion in expenditures has risen by about 0.15 between 2004 and 2014 at the weekly frequency while the cross-sectional dispersion has also gone up by approximately 0.15 over the same period implies that the decreased frequency of expenditures can account for *all* of the rise in inequality of expenditures across households at the weekly frequency of aggregation.

IV. The rise of club stores and expenditure inequality

Previous sections document that U.S. households spend less time shopping and make their shopping trips less frequently so that inequality of *expenditure* measured at high frequency can rise over time while inequality of *consumption* can remain stable. Obviously, there are many possible sources underlying this changing behavior of U.S. households, but one such mechanism is likely the rise of club (warehouse) stores (e.g. Costco, Sam's Club, BJ's) which, by design, sell larger quantities of goods to households at lower unit prices and encourage households to buy goods in bulk. As a result, it has become easier for households to stock up in ways that were not feasible in the past.

Indeed, club stores have expanded dramatically throughout the country since the 1980s (see Panel A of Figure 8), which is consistent with the observed trend in expenditure inequality. To measure intensity of shopping in club stores for a given year, we use the fraction of a household's expenditures that was spent at club stores over the course of that year. Specifically, we calculate the share as $share_{htl}^{(club)} = \sum_{pet} X_{htpl}^{(club)} / \sum_{pet} X_{htpl}$ where $X_{htpl}^{(club)}$ is spending at club

¹² With those same parameter values, we can also verify that $\partial CV_t(X_{h,t}) / \partial CV_h(\bar{X}_h) \approx 1$ so changes in underlying consumption inequality translate one-for-one into changes in expenditure inequality measured at weekly frequency of expenditure aggregation.

stores. In the Nielsen data, the share of household spending at club stores in spending on goods in our sample (food-at-home, alcohol/tobacco, and small nondurables) increased from 7.9 percent (\approx \$320 per year) in 2004 to 9.8 percent (\approx \$390 per year) in 2014 (see Panel B of Figure 8).¹³ Panel C of Figure 8 shows the degree of market penetration of these stores by plotting the distribution of households in the Nielsen data and their distance from the nearest warehouse/club store (in 2004 and 2014). While there is considerable variation in the ease with which households can access one of these retailers, approximately 40 percent of households live less than 5 miles away from one of these stores. At the same time, 30 percent of households have to drive more than 10 miles to reach the nearest store and almost 20 percent have to drive 25 or more miles. Panel D shows the geographical distribution of club stores in 2005.

To assess whether club stores can explain some of the rising concentration in household shopping trips, we characterize the link between how much variation there is in a household's expenditure over time and their reliance on club stores. We do so by regressing households' time-series coefficients of variation on households' club share expenditures, using coefficients of variation measured at different time frequencies. In other words, we estimate the following specification:

$$CV_{htl} = \beta \times share_{htl}^{(club)} + \mathbf{X}_{htl}\boldsymbol{\gamma} + \lambda_t + \psi_h + error \quad (2)$$

where λ_t and ψ_h are the year and household fixed effects, and \mathbf{X} is a vector of controls (the number of children, female head of households, employment status, income brackets, race, employment status of household head, educational attainment of household head, age and age squared for household head). To make inference conservative, we cluster standard errors at the zip-3 level (i.e., first three digits of zip code).

Our theory predicts a positive relationship between time-series CV_{htl} and $share_{htl}^{(club)}$: as a household buys a greater share of their budget at club stores, their purchases should be lumpier. However, causation could run in the opposite direction. For example, if some households choose to have significant time variation in their expenditures (for example, because they like to host a party every month), they might also be more likely to go to club stores to stock up for these events. To rule out this alternative causality, we pursue an instrumental variable approach in which our

¹³ We exclude durables in Nielsen from this analysis since club stores also sell durables—although not in bulk—which would drive up the club share; however, our results are also robust to including durables. In Appendix Table A1, we document which household characteristics are strong predictors of shopping at club stores.

instrument is proximity to a club store (as measured by miles to nearest store). This exploits time-series variation, e.g. stores open and reduce the distance to the nearest club stores faced by some households. To strengthen the quality of our instrumental variables, we exclude households who moved from one location to another. As a result of this restriction, time variation in distance to a club store is determined exclusively by entry/exit of stores.

To construct a measure of distance from club stores, we created a database of geographical locations and openings/closures of club stores for the three largest chains: Sam's Club, Costco, and BJ's. For example, we know that the Costco store in Richmond, CA was opened on October 16th, 1986 at 4801 Central Avenue. A household's distance from the nearest club store is calculated between the centroid of the zip code where a given household lives and the centroid of the zip code of the club store.

The results (Table 3) suggest that shopping at club stores is indeed significantly correlated with more stocking up. First, looking at high frequencies like weekly, there is a positive statistically significant coefficient on the share of expenditures going to club stores, so households who spend relatively more at these stores display more volatility in their expenditures across weeks in a year. However, when we increase the time span over which expenditures are measured, this coefficient shrinks rapidly. At the quarterly frequency, shopping at club stores leads to much less time variation in quarterly spending, which is as expected since it becomes progressively more difficult to stock up for longer periods. Shopping at club stores also explains a diminishing fraction of the variance in households' coefficients of variation at longer durations for time aggregation of expenditures.

Table 3 also shows that the distance to a club store is a strong instrument for the share of spending at club stores in total spending. Households located further from club stores display significantly smaller shares of expenditures at these stores. The first stage F-statistic is above 30. Overall, the OLS and IV estimates are similar. This finding supports the notion that the rising access to club stores has induced households to increasingly stock up on goods and reduce the frequency of their shopping trips. In turn, this change in shopping behavior has generated *spending* patterns that appear more unequal in the cross-section when measured at high frequencies even if their underlying *consumption* flows have not changed.

The results are even stronger if we restrict our attention to households with 2 or more members (Panel B of Table 3), whereas the effects are quite small for single-member households. This is consistent with the idea that club stores are more advantageous for larger households,

whereas single-member households may find it less useful to purchase very large quantities of each type of good. In Appendix Table A1, we document a number of other household characteristics which are associated with higher club store spending, such as education and income. However, we focus on results across all households since we map our estimates into cross-sectional dispersion measures that include all households.

Quantitatively, these estimates are economically significant. From 1980 to 2014, the average expenditure share of club stores has risen by approximately 10 percentage points. Given the IV estimates in Table 3 (Panel A), this implies that club stores can single-handedly account for approximately 40 percent of the trend rise in the cross-sectional dispersion of expenditures measured at the weekly frequency relative to the quarterly or annual frequency, since 0.1×0.375 implies a 0.0375 contribution to the level of dispersion or equivalently 0.0011 per year, relative to an average rise of 0.0029 (see column 1, Panel C, Table 2) over the same time period in the cross-sectional coefficient of variation measured bi-weekly in the CEX for the same set of goods that we use in the Nielsen data.¹⁴ Note that this quantification of the club store contribution may understate the influence of club stores as these stores can influence the behavior of other stores, for example by leading convenience and department stores to start selling multipacks or large packages of goods to keep up with club stores.

V. Conclusion

There has been growing interest in the apparent difference in trend between expenditure and income inequality documented by Slesnick (2001) and Krueger and Perri (2005). Since then, much of the literature has focused on the difficulties associated with measuring expenditure inequality (specifically, under-reporting of expenditures) and concluded that it has, in fact, increased in line with income inequality. We document another measurement issue with consumption measures, namely the infrequent timing of many expenditures, which suggests that consumption inequality has likely increased by *less* than standard measures imply.

Specifically, since households engage in infrequent purchases of many goods, when expenditures are measured at a high frequency many households will appear not to purchase these goods, leading to the appearance of high inequality in consumption, even though their consumption

¹⁴ We can map one-to-one from the effect on time dispersion to the effect on cross-sectional dispersion as shown in section III.B.

may in fact mirror that of households who are observed to purchase the good. We document that households are engaging in fewer shopping trips than in the past and buying larger volumes and quantities when they do make purchases. These trends will, when combined with high-frequency measures of expenditures, lead to the appearance of rising expenditure inequality even when none is present. We show that these patterns can account for much of the rise in expenditure inequality in the Diary Survey of the CEX, and that a lower frequency of aggregation of expenditures points toward little change in consumption inequality. A major force behind this changing consumption behavior appears to be the rise of club/warehouse stores which facilitate and encourage larger sized purchases. As the market for club/warehouse stores becomes more saturated and as bulk goods become more prevalent even in non-club/warehouse stores, one may expect the patterns documented here to have less of an impact on measured spending inequality in the future.

While the increasing prevalence of club retailers appears to have contributed significantly to changing consumer shopping patterns, there are a number of other complementary explanations that could also help account for these trends. For example, anything raising the fixed cost of shopping trips, be it financial (e.g. rising gas prices, rising opportunity cost of time, increasingly moving away from city centers and stores) or in terms of the amount of time (e.g. through rising traffic), would contribute to the declining frequency of shopping. The decline in the real price of gasoline since the early 1980s suggests that gasoline prices are unlikely to have been an important contributor to this changing behavior. However, rising levels of traffic and the progressive “suburbanization” of U.S. cities are more difficult to rule out given the data currently available.

Another force that could lead to a declining frequency of shopping is the cost of storage has declined. For example, increased ownership of refrigerators/freezers has allowed for more storage of food products, but this is unlikely to be an important contributor since we observe increased stocking up across a wide range of goods, not just food products. The growing size of U.S. houses, on the other hand, could induce more stocking up on the part on households. In the absence of detailed information on the changing sizes of homes across regions, it is difficult to quantify this channel precisely with our data although this would be a promising area for future research.

A final force worth considering is financial innovation. Much of the work focusing on consumption inequality has explained the flat profile found in the Interview Survey through the financial innovation channel. According to the leading hypothesis, expanded access to credit has allowed households to better smooth transitory economic shocks, thereby pushing down

consumption inequality, even though the prevalence of transitory shocks, reflected in rising income inequality, has been increasing. This same financial innovation expanding credit may also have allowed households to better take advantage of bulk discounts like those available at club/warehouse stores. Interestingly, the expansion of credit may have acted to raise spending inequality through our mechanism when looking at high-frequency shopping patterns, while reducing this inequality in lower frequency data by allowing households to mitigate transitory income shocks. In future work, we intend to use data as in Gelman et al. (2016) to examine the relationship between the availability of credit and bulk shopping.

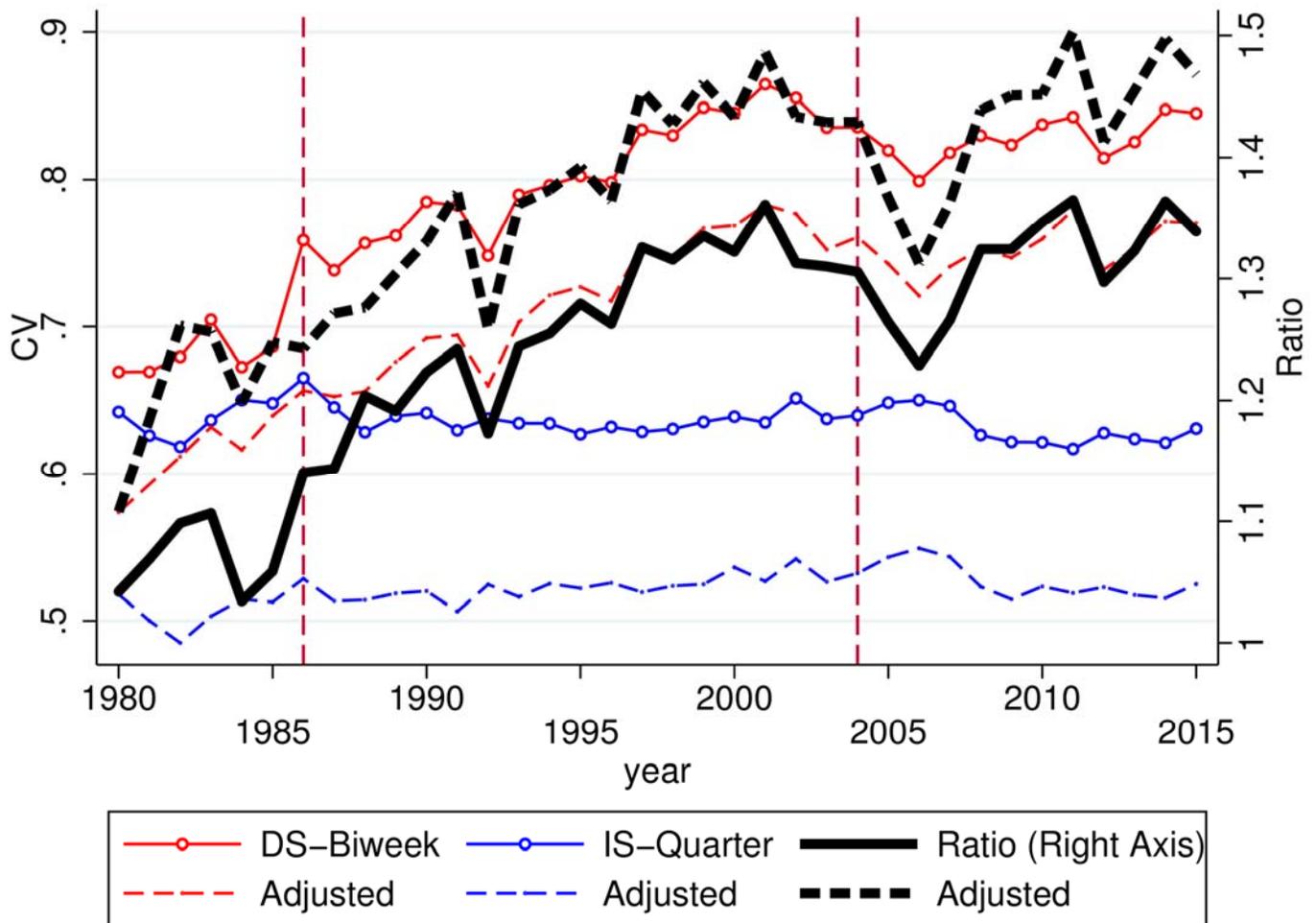
While many forces have therefore likely contributed to a reduced frequency of shopping, the growing prevalence of online retailing and home deliveries is reducing fixed costs of shopping and should therefore be pushing toward a *higher* frequency of shopping. As online retailing continues to grow in both size and scope, this implies we may observe a reversal of some of the patterns documented here. It therefore seems promising for future work to consider how these different forces will balance out so that we can better understand how to properly measure underlying trends in consumption inequality.

References

- Aguiar, Mark, Erik Hurst, and Loukas Karabarbounis 2013. "Time Use during the Great Recession," *American Economic Review* 103(5), 1664-1696.
- Aguiar, Mark and Erik Hurst, 2013. "Deconstructing Lifecycle Expenditure," *Journal of Political Economy* 121(1), 437-92.
- Aguiar, Mark, and Mark Bilal, 2015. "Has Consumption Inequality Mirrored Income Inequality?" *American Economic Review* 105(9), 2725-56.
- Attanasio, Orazio, Erich Battistin, and Hidehiko Ichimura, 2007. "What Really Happened to consumption Inequality in the United States?" in *Hard-to-Measure Goods and Services: Essays in Honor of Zvi Griliches*, Ernst R. Berndt and Charles R. Hulten, eds., University of Chicago Press.
- Attanasio, Orazio, Erik Hurst, and Luigi Pistaferri, 2015. "The Evolution of Income, Consumption, and Leisure Inequality in the United States, 1980-2010," in *Improving the Measurement of Consumer Expenditures*, Christopher Carroll, Thomas Crossley, and John Sabelhaus, editors, p. 100 – 140, University of Chicago Press.
- Attanasio, Orazio P. and Luigi Pistaferri, 2016. "Consumption Inequality." Forthcoming in the *Journal of Economic Perspectives*.
- Baker, Scott, and Lorenz Kueng, 2017. "Shopping for Lower Sales Tax Rates," manuscript.
- Battistin, Erich, 2003. "Errors in survey reports of consumption expenditures," IFS Working Papers W03/07, Institute for Fiscal Studies.

- Bee, Adam, Bruce D. Meyer, and James X. Sullivan. 2012. "The Validity of Consumption Data: Are the Consumer Expenditure Interview and Diary Surveys Informative?" NBER Working Paper No. 18308.
- Bee, Adam, Bruce D. Meyer, and James X. Sullivan. 2015. "The Validity of Consumption Data: Are the Consumer Expenditure Interview and Diary Surveys Informative?" in *Improving the Measurement of Consumer Expenditures* (2015), Christopher D. Carroll, Thomas F. Crossley, and John Sabelhaus, editors (p. 204 - 240).
- Blundell, Richard, Luigi Pistaferri and Ian Preston, 2008. "Consumption Inequality and Partial Insurance." *American Economic Review* 98(5), 1887-1921.
- Coibion, Olivier, Yuriy Gorodnichenko, Lorenz Kueng, and John Silvia, 2012. "Innocent Bystanders? Monetary Policy and Inequality in the U.S." NBER Working Paper 18170.
- Gelman, Michael, Yuriy Gorodnichenko, Shachar Kariv, Dmitri Koustas, Matthew D. Shapiro, Dan Silverman, Steven Tadelis, 2016. "The Response of Consumer Spending to Changes in Gasoline Prices," NBER Working Paper 22969.
- Gorodnichenko, Yuriy, Klara Sabiniarova Peter, and Dmitriy Stolyarov, 2009. "Inequality and Volatility Moderation in Russia: Evidence from Micro Level Panel Data on Consumption and Income," *Review of Economic Dynamics*, 13(1), 209-237.
- Krueger, Dirk and Fabrizio Perri. 2005. "Does Income Inequality Lead to Consumption Inequality? Evidence and Theory." *Review of Economic Studies*, 73(1) 163-193.
- Krueger, Dirk, Fabrizio Perri, Luigi Pistaferri, and Giovanni L. Violante, 2010. "Cross sectional Facts for Macroeconomists," *Review of Economic Dynamics* 13(1), 1-14.
- Lee, Jungmin, Daiji Kawaguchi, and Daniel S. Hamermesh. 2012. "Aggregate Impacts of a Gift of Time." *American Economic Review* 102 (3): 612–16.
- Menzio, Guido and Nicholas Trachter, 2015. "Equilibrium Price Dispersion with Sequential Search," *Journal of Economic Theory*, 160 (6), 188-215.
- Michaillat, Pascal and Emmanuel Saez. 2015. "Aggregate Demand, Idle Time, and Unemployment," *Quarterly Journal of Economics* 130(2), 507-569.
- Piketty, Thomas and Emmanuel Saez, 2003. "Income Inequality in the United States, 1913-1998." *Quarterly Journal of Economics*, 118(1), 1-39.
- Piketty, Thomas, Emmanuel Saez, and Gabriel Zucman, 2016. "Wealth Inequality in the United States since 1913, Evidence from Capitalized Income Tax Data." *Quarterly Journal of Economics* 131(2), 519-578.
- Slesnick, Daniel T. 2001. Consumption and Social Welfare. Cambridge: Cambridge University Press.
- Wong, Arlene. 2016. "Population Aging and the Transmission of Monetary Policy to Consumption," manuscript.

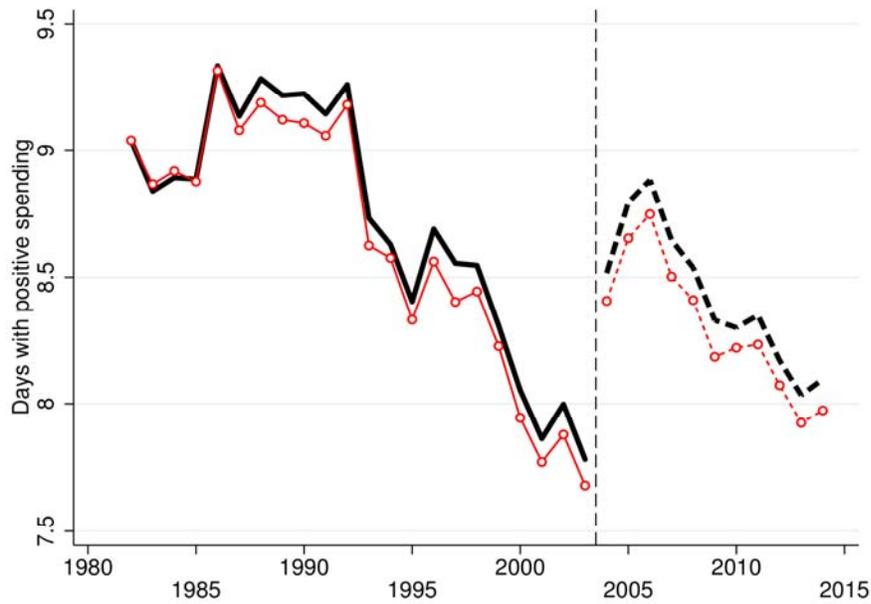
Figure 1. Spending inequality in CEX Diary Survey and CEX Interview Survey.



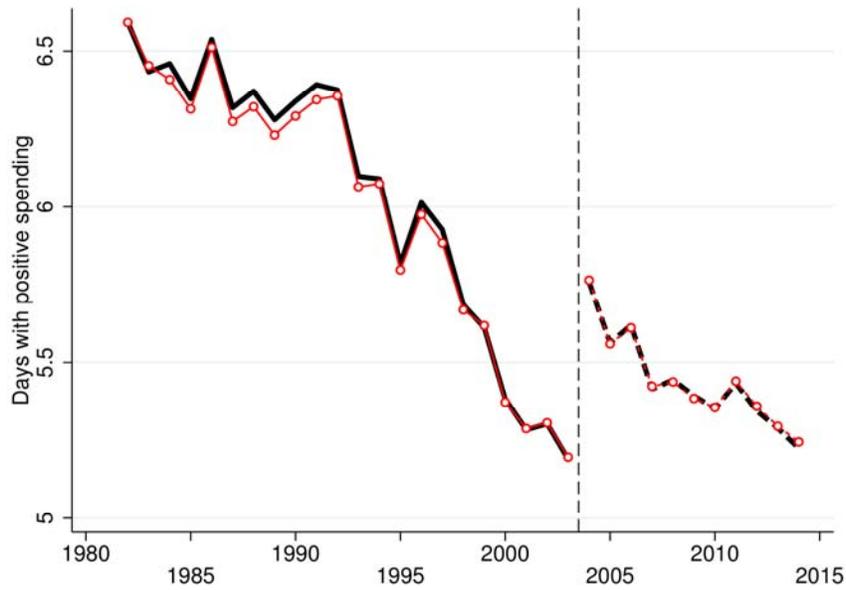
Notes: The figure plots the coefficient of variation (CV on left axis) of expenditures on non-durable goods and services across households in the Diary survey (DS-biweekly) and Interview survey (IS-quarterly) over time. See section 1 for more details on the construction of these measures. The ratio of the two DS/IS is plotted using the bold black line and measured on the right axis. Solid lines are raw measures while dashed lines are residual measures, as described in section 1. Vertical lines indicate major structural breaks in diary survey design.

Figure 2. Frequency of shopping, CEX Diary Survey.

Panel A. All Nondurables

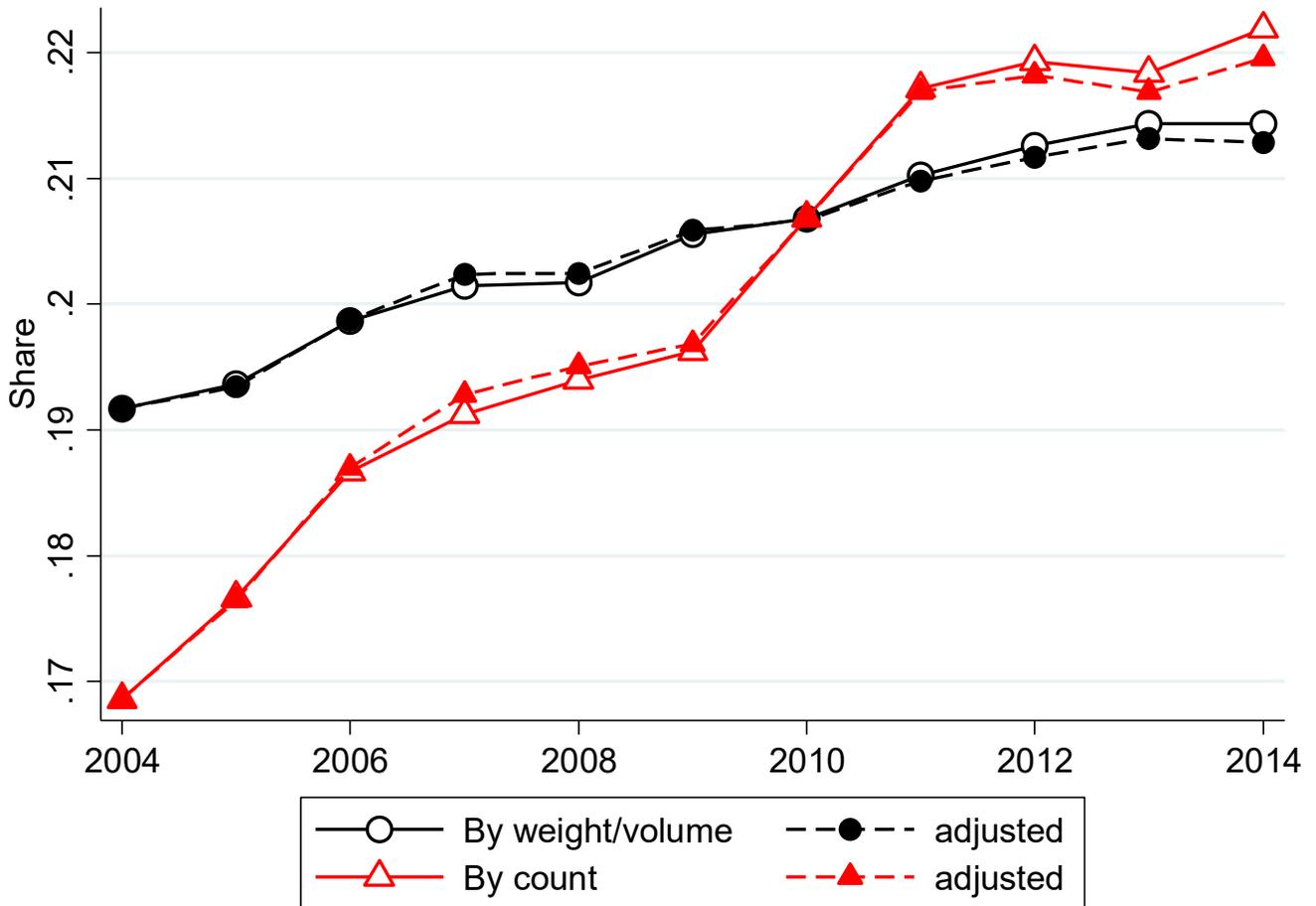


Panel B. Food at Home, Alcohol/Tobacco, and Small Nondurables



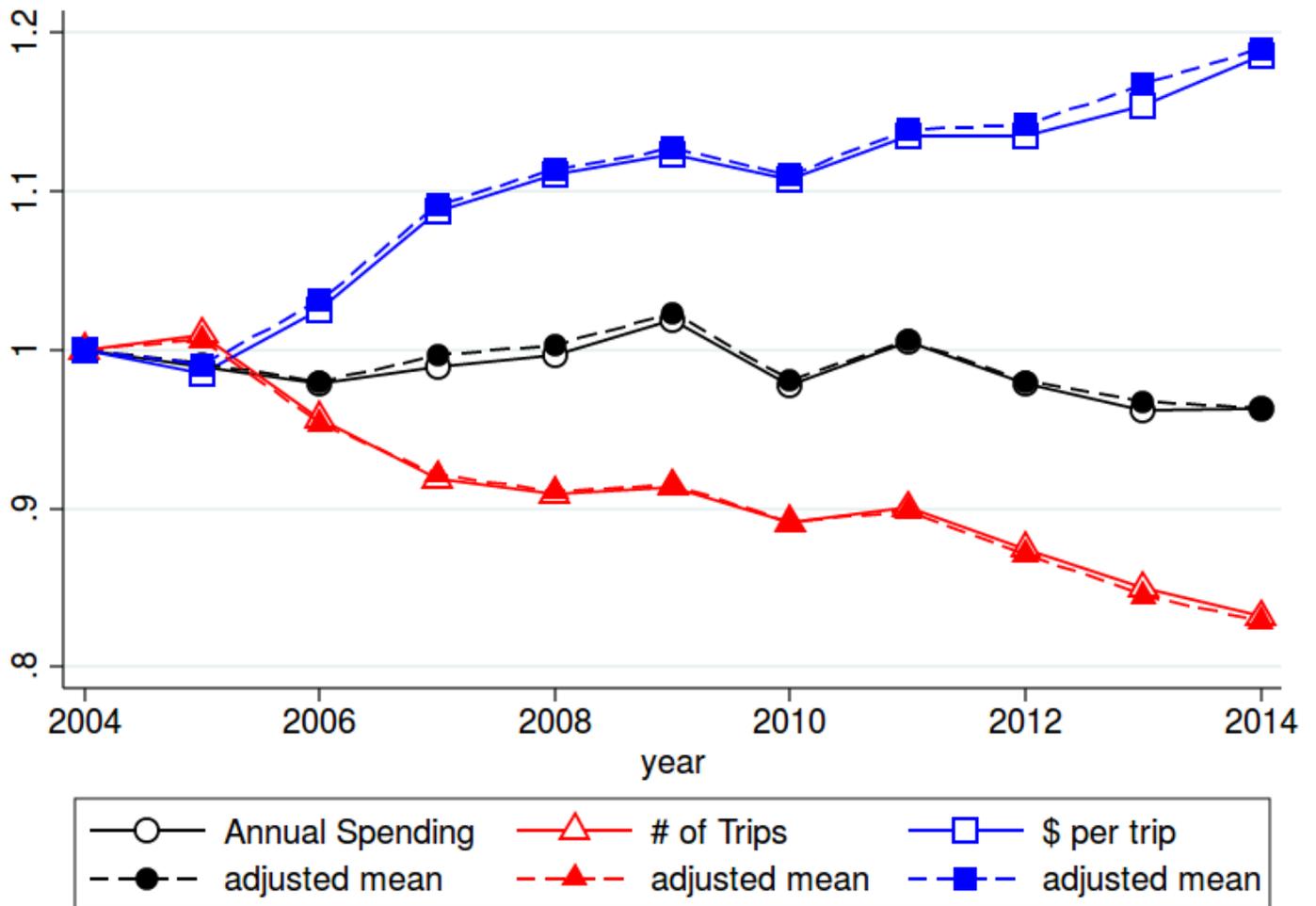
Notes: The figure plots the average number of days in which households report any positive spending in CEX (measured for 2 week periods) over time. The vertical dashed line shows the time when the CEX Diary Survey had a change in how it collects data. See section 1 for more details. In Panel A, the sample is all nondurables (see Appendix B). In Panel B, the set of goods is restricted to be common to both the CEX surveys and the Nielsen data. The included categories are food-at-home, alcohol/tobacco, and small non-durables which matches the coverage of goods in our Nielsen sample.

Figure 3. Share of large-volume purchases.



Notes: The figure shows the dynamics of the share of large-volume purchases in total purchases. Large-volume purchases are identified as purchases that exceed the 90th percentile of the distribution of the purchased weights or counts in 2004. Expenditure shares are used to weigh product modules. Sampling weights are used to aggregate across households. Solid lines with empty markers show the dynamics of the raw averages. Dashed lines with filled markers show the dynamics adjusted for changes in household characteristics (quadratic polynomial in the age of household head's age and a set of dummy variables for household size, employment status of household head and his/her spouse, number of children, and race). Approximately 55% of universal product codes (UPCs) are measured in ounces and 45% are measured in counts. See section 2 for details.

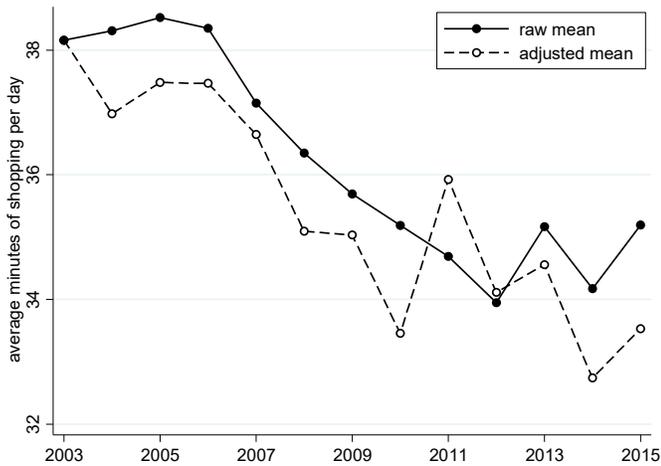
Figure 4. Consumer spending, number of shopping trips, and spending per trip in ACNielsen household panel.



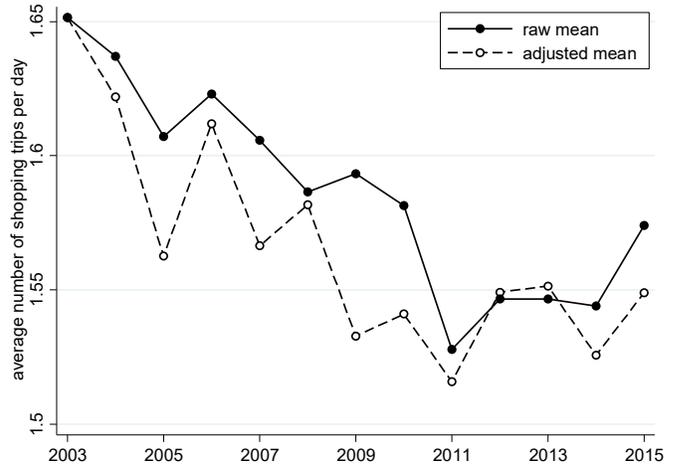
Notes: Solid lines with empty markers show the dynamics of the raw averages. Dashed lines with filled markers show the dynamics adjusted for changes in household characteristics (quadratic polynomial in the age of household head's age and a set of dummy variables for household size, employment status of household head and his/her spouse, number of children, and race). The black lines are the average log spending per year. The red lines are the average number of trips per year. The number of trips is the number of trips where the household scanned at least one UPC barcode. The blue lines are the average log spending per shopping trip in a given year. All series are normalized to one in year 2004. Spending is adjusted for inflation using the "Personal Consumption Expenditures (PCE): Chain-type Price Index" (FRED Series: PCEPI). See section 2 for details.

Figure 5. Shopping time.

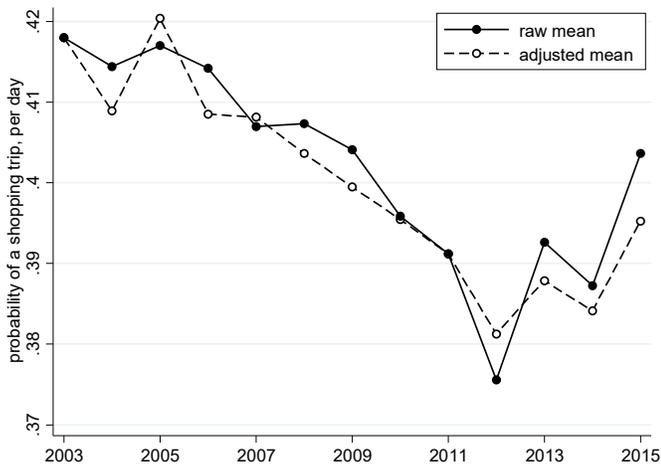
Panel A: Shopping time for purchases of goods.



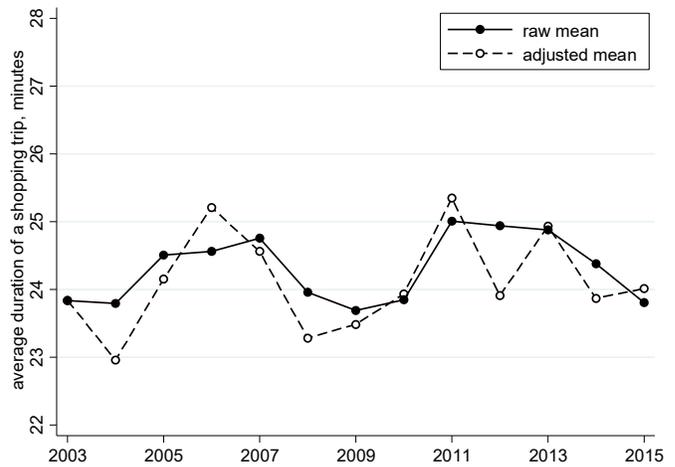
Panel B. Number of trips per day (conditional on having a trip) for purchases of goods



Panel C: Probability of a shopping trip for purchases of goods



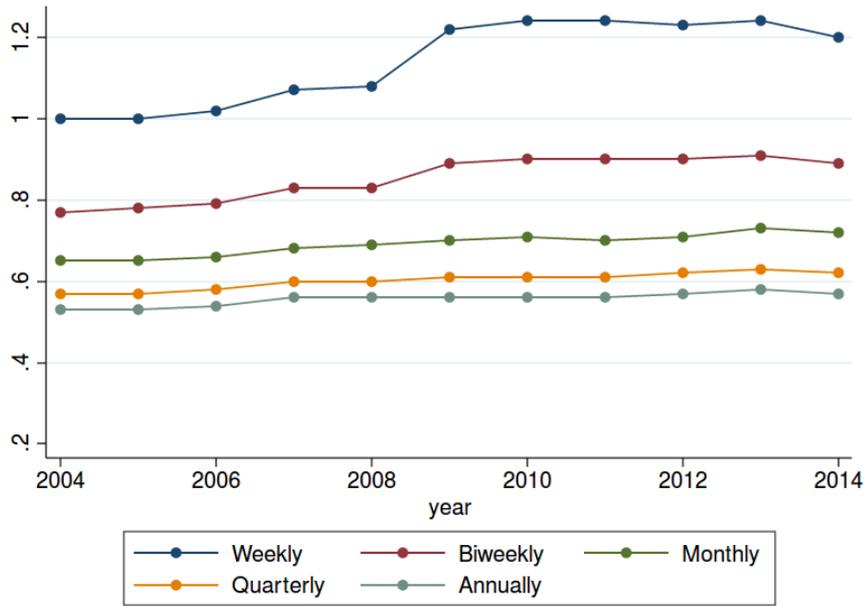
Panel D: Average duration of a shopping trip for purchases of goods



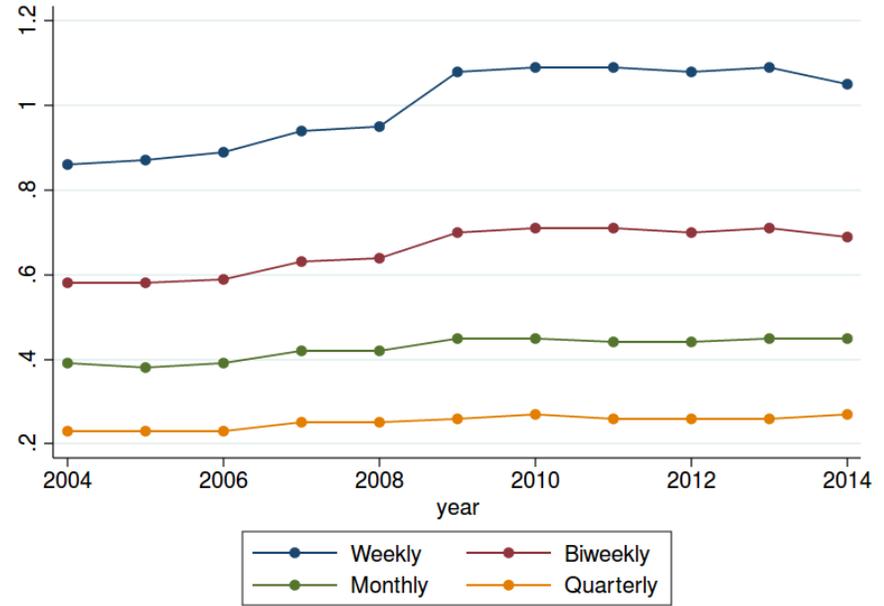
Notes: Panel A reports total shopping time (includes travel and other purchase related activities). Panel B reports the number of shopping trips per day conditional on having a shopping trip. Panel C reports the probability of having a shopping trip on a given day. Panel D reports the average duration of a shopping trip (including travel time and other purchase related activities; conditional on having a shopping trip). The black, solid line shows the raw average. The black, dashed line shows the average (regression) adjusted to demographic changes. See section 2 for details.

Figure 6. Inequality in spending by frequency of time aggregation.

Panel A: Cross-Sectional Dispersion of Expenditures

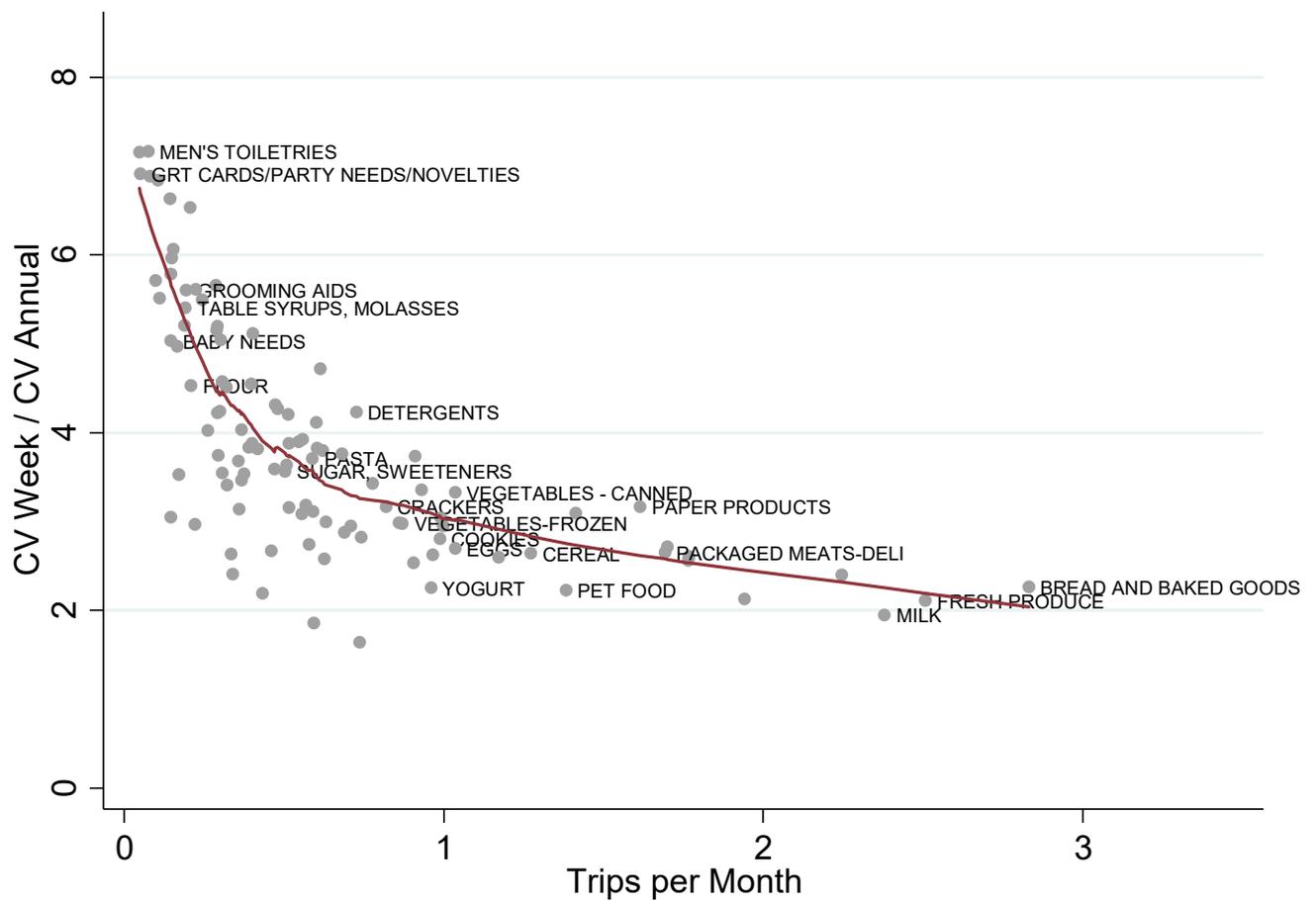


Panel B: Average Time-Series Dispersion of Expenditures



Notes: The figures plot coefficients of variation (CV) of household expenditures when average expenditures are measured at different frequencies of time aggregation ranging from weekly to annually. The left panel shows the average annual cross-sectional coefficient of variation of expenditures across households, where expenditures are measured at different time frequencies. The right panel shows the average (across households) time-series coefficient of variation of each household's expenditures over the course of the year, measuring expenditures at different frequencies. All calculations are for the Nielsen data. See section 3 for details. The corresponding figure for inequality after controlling for household characteristics are in Appendix Figure A7.

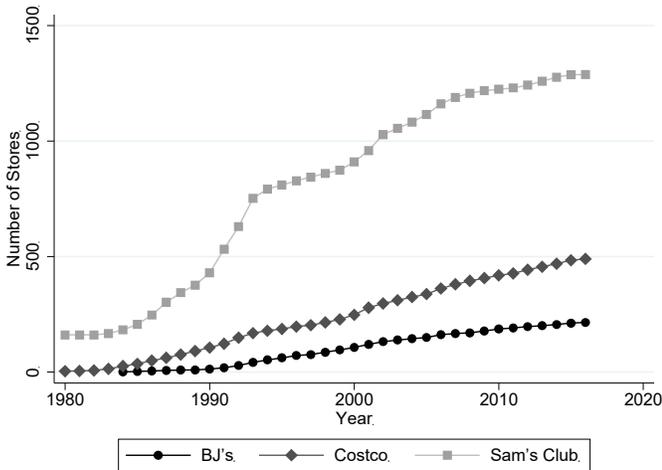
Figure 7. Cross-sectional inequality and the frequency of purchases



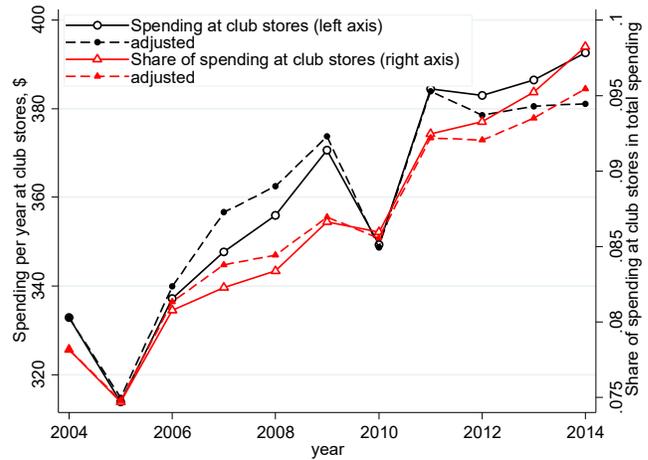
Notes: The horizontal axis shows the average number of shopping trips per month for a given module of goods in the Nielsen data. These data are from Baker and Kueng (2017). The vertical axis shows the ratio of the coefficient of variation at the weekly frequency to the coefficient of variation at the annual frequency for a given module for year 2014. Each point in the scatter plot corresponds to a module. The solid, red curve shows fitted values from the locally weighted regression (lowess).

Figure 8. Importance of club stores

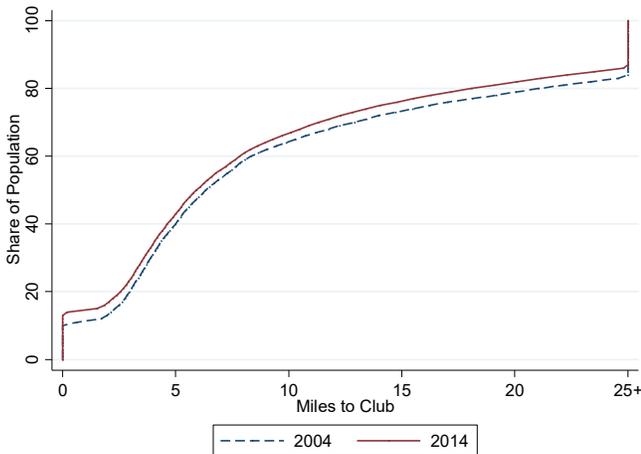
Panel A. Penetration of club stores.



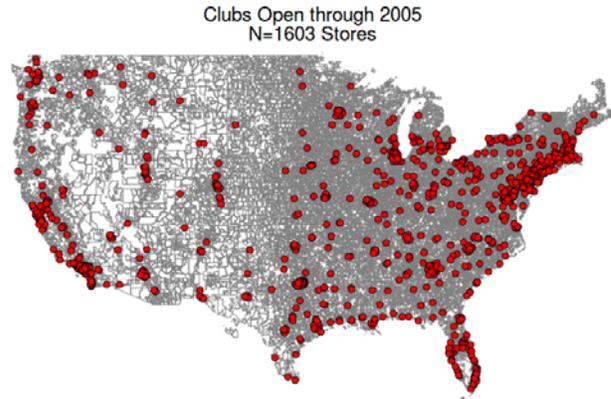
Panel B. Spending in club stores



Panel C. Distance to nearest club store



Panel D. Geographic distribution of club stores



Notes: Panel A plots the numbers of different club/warehouse stores over time. Panel B plots the average dollar amount of spending per household at club/warehouse stores and the average share of expenditures by households at these retailers for the goods covered in our sample (food-at-home, alcohol/tobacco, and small non-durables). Panel C plots the distribution of distances from the nearest club/warehouse retailer for households in Nielsen sample in 2004 and 2014. Panel D shows the distribution of club stores in our sample in 2005. See section 4 for details.

Table 1. Biweekly Spending in the CEX Diary Survey and Nielsen data

Spending category	CEX Diary				Nielsen			
	Mean	St.Dev.	IQR	Zero Share	Mean	St.Dev.	IQR	Zero Share
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
TOTAL SPENDING	1,475.37	1,576.94	1,519.93	0.00	161.16	125.04	151.00	0.10
SELECTED COMPARABLE CATEGORIES OF NONDURABLE GOODS								
Baby food	19.41	27.90	11.40	0.95	19.95	27.94	16.28	0.98
Pet food	25.36	28.09	22.97	0.72	19.86	21.21	20.70	0.68
Cereal	7.32	5.84	5.82	0.53	9.01	7.87	8.01	0.53
Coffee	11.14	9.23	7.64	0.69	11.67	10.93	8.98	0.77
Crackers	5.11	3.66	3.87	0.67	4.59	3.62	3.71	0.72
Eggs	4.49	3.20	3.20	0.50	3.55	2.46	2.17	0.63
Milk	7.81	6.32	6.51	0.27	6.63	5.35	5.16	0.37
Fresh meat	32.52	30.63	31.52	0.25	10.05	7.61	7.63	0.88
Detergent	10.23	8.49	9.63	0.69	7.95	7.15	7.78	0.75
Beer	26.32	23.24	21.61	0.83	24.05	23.73	19.24	0.91
Liquor	28.89	24.90	24.21	0.96	27.94	28.35	25.95	0.93
Fresh produce	23.82	22.39	24.50	0.17	9.19	9.00	8.91	0.41
Lawn and garden	29.65	51.24	20.49	0.89	11.79	12.59	10.48	0.94
Hair care products	13.06	16.65	9.51	0.79	8.04	7.76	6.93	0.82
Over the counter drugs	14.99	15.52	13.21	0.74	14.97	16.28	15.51	0.54
Oral hygiene	7.32	6.30	5.91	0.80	6.26	6.46	5.36	0.78
Shaving needs	11.47	10.91	10.36	0.94	8.55	9.53	8.40	0.93
Vitamins	24.83	33.05	17.15	0.91	19.26	19.48	17.91	0.79
MATCHED NONDURABLES	239.20	181.94	219.81	0.07	151.47	117.47	143.00	0.10
Biweekly Observations	6,241				1,199,031			

Notes: Columns (1) and (5) show the mean of spending in the CEX Diary Survey and AC Nielsen, respectively, conditional on making a purchase, over a biweekly period in 2014. Columns (2) and (6) show the standard deviation of this spending across households. Columns (3) and (7) show the interquartile range (IQR) of this spending across households. Columns (4) and (8) show the zero share of spending on the specified category in the biweekly period in 2014. For the CEX Diary survey, the sample of households is restricted to households reporting two diary weeks. By construction, the Diary Survey has no household with 0 spending in the biweekly period. For AC Nielsen, the sample of households includes only households with at least one shopping trip in each month of 2014. We aggregate daily spending to the biweekly period (weeks 1 and 2 of 2014 are one biweekly period, weeks 3 and 4 are a biweekly period, etc.) and treat the data as repeated cross-sections when calculating moments.

Table 2. Time trends in expenditure inequality by time aggregation.

Dep. var.: Coefficient of variation	Frequency of aggregation				
	Weekly (1)	Biweekly (2)	Monthly (3)	Quarterly (4)	Annual (5)
Panel A: Nielsen data, 2004-2014.					
Year	0.0272*** (0.005)	0.0154*** (0.0025)	0.0078*** (0.0008)	0.0058*** (0.0006)	0.0046*** (0.00039)
Observations	11	11	11	11	11
Panel B: CEX data (all nondurables), 1980-2015.					
Year	0.0056*** (0.0011)	0.0048*** (0.0009)		-0.0004 (0.0003)	-0.0001 (0.0004)
Observations	36	36		36	36
Panel C: CEX data (nondurables as in the Nielsen data), 1980-2015.					
Year	0.0028*** (0.0004)	0.0020*** (0.0004)		-0.0013** (0.0002)	-0.0012** (0.0002)
Observations	36	36		36	36

Notes: the table reports estimated slope in the regression of coefficient of variation for a given frequency of time aggregation on time trend. Time aggregation is indicated in the top row. Panel A uses data from AC Nielsen. Panel B uses CEX data covering all non-durable goods and services: the Diary Survey for columns (1) and (2) and the Interview Survey of columns (4) and (5). For the Interview Survey of the CEX, the dependent variable in column (4) includes some expenditures that are measured at the monthly frequency. Panel C restrict the CEX data to cover only goods included in the Nielsen sample (food-at-home, alcohol/tobacco, and small non-durables). Newey-West standard errors are reported in parentheses. ***,**,* denote statistical significance at 1, 5 and 10 percent levels.

Table 3. Lumpiness of purchases and shopping at club stores

Panel A. Full Sample

Dep. var.: Coefficient of variation	Frequency of aggregation							
	Weekly		Biweekly		Monthly		Quarterly	
	OLS (1)	IV (2)	OLS (3)	IV (4)	OLS (5)	IV (6)	OLS (7)	IV (8)
Club share	0.238*** (0.012)	0.376* (0.222)	0.148*** (0.009)	0.220 (0.165)	0.066*** (0.006)	-0.019 (0.128)	0.012*** (0.006)	-0.059 (0.093)
N	393,822	393,822	393,822	393,822	393,822	393,822	393,822	393,822
R2	0.768	0.768	0.719	0.719	0.623	0.622	0.482	0.481
1 st stage <i>F</i> -stat		38.14		38.14		38.14		38.14

Panel B. Families

Dep. var.: Coefficient of variation	Frequency of aggregation							
	Weekly		Biweekly		Monthly		Quarterly	
	OLS (1)	IV (2)	OLS (3)	IV (4)	OLS (5)	IV (6)	OLS (7)	IV (8)
Club share	0.224*** (0.014)	0.569*** (0.215)	0.136*** (0.011)	0.348** (0.163)	0.057*** (0.008)	0.052 (0.128)	0.011*** (0.006)	0.004 (0.096)
N	290,742	290,7442	290,742	290,742	290,742	290,742	290,742	290,742
R2	0.778	0.775	0.732	0.729	0.637	0.637	0.503	0.503
1 st stage <i>F</i> -stat		37.45		37.45		37.45		37.45

Notes: The dependent variable is the coefficient of variation (CV) calculated as follows. For each household, we calculate i) standard deviation of spending at a given frequency (weekly, biweekly, monthly, quarterly) for a given year and ii) average spending per period (total annual spending divided by the number of periods with shopping trips). The coefficient of variation (CV) is i) divided by ii) so that CV is time-series volatility of spending for a given household in a given year. *Club share* is the share of annual spending at club stores (Sam's Club, Costco, BJ's, etc.) in total annual spending at all stores. Spending includes only food, alcohol/tobacco, and small nondurables (paper towels, toothpaste, etc.). The sample of households includes only households with at least one shopping trip in each month of a given year. For each household, the instrumental variable is the distance to the closest club store (Sam's Club, Costco, BJ's). This distance is calculated between the centroid of the zip code where a given household lives and the centroid of the zip code where the nearest club store is located. Regressions include but do not report coefficients on the following controls: year and household fixed effects, age and age squared for the household head, a set of dummy variables for household income brackets, number of children, employment status, race, educational attainment, gender of household head. Standard errors are clustered at the zip-3 level (i.e., first three digits of zip code). ***, **, * denote significance at 1, 5, and 10 percent levels. Panel A is for all households, while Panel B restricts to households with a household size of at least two persons.