Revisiting the Consumption-Retirement Puzzle

Niklas Flamang*†

November 18, 2022

Updated frequently. Click here for most recent version.

Abstract

Is the post-retirement expenditure drop a true drop in consumption? Leveraging a long panel data set, I construct a novel measure of consumption that decomposes post-retirement expenditure declines into: 1) savings effects through lower barcode-level prices and bulk purchases, 2) quality effects from substituting towards cheaper products, and 3) quantity changes. Estimating event studies around retirement, I find that paying lower prices for identical products explains at most 8.5% of the expenditure drop. The majority of the expenditure drop results from declining quantities and declining quality of the consumption bundle. I further show that home production cannot explain the expenditure drop as spending on home production inputs also falls. I also study shopping adjustments when households become unemployed and find similar results.

*University of California, Berkeley. I am grateful to my advisors Emi Nakamura, Yuriy Gorodnichenko, and Amir Kermani for their invaluable guidance throughout this project. For helpful discussions and valuable suggestions, I am thankful to Hadar Avivi, Stefano DellaVigna, Sree Kancherla, Pat Kline, Quan Le, Michael Love, Enrico Moretti, Adam Rosenberg, Benjamin Schoefer, Jón Steinsson, Damian Vergara, Chris Walters, Harry Wheeler, Danny Yagan as well as participants of the Berkeley Macro Seminar, the Berkeley Graduate Economics Mini Symposium, the Berkeley Macro Lunch, the Berkeley Finance Lunch, and the Berkeley Public Economics Seminar. All errors are my own.

†Researcher(s)' own analyses calculated (or derived) based in part on data from Nielsen Consumer LLC and marketing databases provided through the NielsenIQ Datasets at the Kilts Center for Marketing Data Center at The University of Chicago Booth School of Business. The conclusions drawn from the NielsenIQ data are those of the researcher(s) and do not reflect the views of NielsenIQ. NielsenIQ is not responsible for, had no role in, and was not involved in analyzing and preparing the results reported herein.

1 Introduction

Every day, over 10,000 people in the United States reach retirement age. Given the large size of the Baby Boomer generation, this pattern will continue for at least another fifteen years. The aging population has wide-reaching implications for the fiscal position of the United States, its future growth trajectory, and the welfare of the soon-to-be retirees themselves. This latter effect is a source of immediate concern especially because Covid-19 has pushed many people into retirement early. This wave of retirements may have have big welfare consequences as consumer expenditures decline sharply in the first years of retirement (the "post-retirement consumption drop"). Whether this decline is a measurement artifact (e.g., retirees can search for and pay lower prices thus driving a wedge between consumption and nominal spending) or a genuine decrease in consumption still is a matter of debate. Using very rich scanner data covering sixteen years, I document that price effects are relatively small and thus a significant decrease in consumption is likely real.

In this paper, I combine evidence from rich scanner survey data to show that expenditures fall starkly at retirement. In the first six years of retirement, total non-durable expenditures fall by a total of thirteen percent while scanner-data covered expenditures (which mostly spans food at home broadly defined) fall by around seven in the same time frame. I then leverage the scanner data to decompose household expenditure patterns. To do so, I define three margins of shopping behavior: price effects, bulk effects, and quality adjustments. Price effects capture purchasing a given barcode-level item at a lower price. This can be achieved through buying items on sale, using coupons or traveling to cheaper stores. Bulk effects are savings that arise from a households' allocation of purchases to bulk quantities (and thus lower per-unit prices). Quality adjustments represent moves along the quality ladder for very similar items (e.g., buying pasture-raised eggs vs. battery-cage eggs). Equipped with these measures of shopping adjustment, I re-estimate the expenditure drop after correcting for price and bulk effects. Three years into retirement, price savings and bulk savings account for a 0.25 percentage point drop in expenditure while uncorrected expenditures have fallen by 2.85 percent. Six years into retirement, some of these savings have faded away so that price and bulk savings contribute 0.2 percentage points to the total expenditure drop of 7.25%. Together, price, bulk, and quality effects contribute 1 percentage point to the expenditure drop at three years and 1.25 percentage points after six years. Nevertheless, even correcting expenditures for all of these margins, six percentage points of the expenditure drop after six years are the result of changes in *quantities*.

These expenditure and quantity declines are not explained by either work related expenditures or substitution towards home production: Expenditures on goods unrelated to work fall and expenditures on home production inputs also decline, a result that holds both within scanner and survey data. In addition, there is little evidence of expenditure timing: Durable expenditures fall by substantially more than non-durable expenditures while there is no evidence of large durable purchases right before retirement. All of these results point towards the post-retirement expenditure drop representing a true decline in consumption.

To investigate how these patterns vary across households, I perform a variety of tests. Expenditure and quality drops are smallest for those households that see the smallest post-retirement income drop, pointing towards a significant role for current income in explaining consumption dynamics. Similarly, households with higher wealth are more insulated from the post-retirement income drop and have smaller expenditure drops than households with lower levels of saving, particularly with respect to non-durables. For price and bulk purchasing adjustments, the patterns are less stark. Price adjustments are largest for households with the largest post-retirement income drop, but beyond this group it is not clear if prices fall the most for households for whom expenditures decline the most. Consequently, raw expenditure adjustments look very similar to price-and-bulk corrected expenditure adjustments for most households.

Analyzing whether shopping effects are important in other contexts, I find that households with female heads not working full time pay substantially lower prices, but that much of that effect seems to be driven by fixed differences across households rather than within-household variation over time. Tracing out shopping behavior over the course of an non-employment spell, I find that similar to the case of retirement, the primary margin of shopping adjustment is the quality of the consumption bundle rather than purchasing identical products at lower prices. These results suggest that expenditures are generally a good proxy for consumption expenditures.

My findings have important implications: The large drop of consumption at the onset of retirement is not only inconsistent with the life-cycle permanent income hypothesis, it is also not explained by consumption models that incorporate credit market frictions and uninsurable income risk. While credit frictions can explain why households cannot smooth consumption early in life, the stark drop in consumption at retirement is inconsistent with consumption smoothing. In addition, the consumption drop at retirement is not explained by consumption models with little holdings of liquid assets that result from return differentials between high-return illiquid and low-return liquid assets: First, many illiquid assets become liquid at retirement (e.g., 401k plans and IRAs, pension plans). Second, the decline of *current* income at retirement is permanent, making it hard to justify based on one-time adjustment costs of changing the asset allocation. However, consumption adjustments are smaller for wealthier households, suggesting an important role of differences in wealth accumulation in explaining the drop. Consistent with prior results in the literature, households who see the smallest post-retirement income drop also see consumption fall the least, implying the need for consumption models that can generate sensitivity to current income beyond rational illiquidity or credit market frictions.

Beyond our understanding of consumption behavior, my findings also have important welfare implications. With the baby boomers starting to enter retirement, the United States economy will see four to five million retirements every year over the next fifteen years. To the extent that these retirements go along with large declines in expenditures, this pattern is likely to present a future drag on domestic consumption and hence total output growth. The apparently large sensitivity to current income also suggests that the optimal design of old-age retirement schemes should consider disbursement policies as an important policy lever. Many households do not annuitize their security holdings and they are also unwilling to "eat their house" as evidenced by the relative dearth of reverse mortgage products. Designing private retirement accounts so that asset holdings are easily transformed into constant cash flows seems like a promising way of preventing excessively large expenditure declines in retirement. More generally, my findings further suggests that retirement schemes that guarantee constant cash flows may have substantial welfare benefits relative to systems without recurring payouts.

The main contribution of this paper is to investigate to what extent the post-retirement expenditure drop is a result of measurement error or if it is a real drop in consumption. Starting in the 1980s, research has documented that households appear to under-save for retirement resulting in substantial declines in consumption in old age (Hamermesh 1984, Bernheim 1987, Hausman and Paquette 1987). This finding has sparked a wave of interest in explanations of the post-retirement expenditure drop that are consistent with the life-cycle permanent income hypothesis. These explanations are broadly categorizable into three groups: First, households' may have a lower marginal utility of consumption later in life, for example resulting from changes in household composition and aging itself. Since optimal smoothing of life-cycle consumption predicts the equalization of *marginal utilities*, not consumption levels, this may explain why expenditures drop so much later in life (Banks et al. 1998). Second, changes in expenditures might be explained by declines in work-related expenses. If the post-retirement expenditure drop is driven by these, then utility-relevant consumption has not fallen at all (Hurd and Rohwedder 2005). Third, retired households may exert shopping effort to lower the prices they face and engage in more home production so that lower expenditures translate into constant consumption levels (Aguiar and Hurst 2005, Aguiar and Hurst 2007, Hurst 2008).

Notably, neither of the first two explanations seem sufficient to explain the retirement savings puzzle. Banks et al. (1998) show that expected changes in household composition and mortality are not sufficient to equalize pre-retirement and post-retirement marginal utilities of consumption in British micro data. In addition, while work-related expenses in their data are responsible for a large share of the post-retirement expenditure drop, expenditure in *all* categories of consumption falls. More recently, Olafsson and Pagel (2018) use data from a personal finance aggregator to show that spending in both leisure and work-related categories of consumption falls while savings actually *increase* in retirement. Last, Stephens and Toohey (2018) leverage forty years of cross-sectional data as well as longitudinal evidence to show that caloric intake falls during retirement.

In addition to the literature on the retirement-savings puzzle, this paper contributes to the literature on household-level prices and the adjustment of shopping behavior in light of shocks. Aguiar and Hurst (2007) were the first to utilize scanner data to construct household-level price indices. In terms of deconstructing deviations from the price of the average bundle, this paper is closely related to Kaplan and Menzio (2015) and Nevo and Wong (2019). Kaplan and Menzio (2015) use the 2004-09 years of the KNCP to decompose prices into a store component, a store-specific goods component, and a transaction component plus three covariance terms. They find that around 90% of the price dispersion across households comes from the store and store-good components, suggesting that at the household level, the choice of retailer is what is driving price dispersion. In related work, Coibion et al. (2015) show that inflation in *effective* prices paid by consumers declines significantly with higher unemployment rates while posted prices remain relatively unchanged. These effects are mostly driven by households switching between retailers rather than purchasing on-sale items. Another closely related paper is Nevo and Wong (2019) who use scanner data to investigate to what extent costly shopping activities lower household-level prices. Here, too, the household-level price index is defined as a cross-sectional measure that compares the cost of the household's bundle at actual prices to the cost of the same bundle at average prices.

An important difference between this paper and both Kaplan and Menzio (2015) and Nevo and Wong (2019) is that both Kaplan and Menzio (2015) and Nevo and Wong (2019) are agnostic about the "correct" level of aggregation for the price index, allowing or increasingly broad comparisons across products. Meanwhile, I follow Aguiar and Hurst (2007) and think of the household-level price index as a measure of prices paid for identical goods. This allows me to interpret the choice of exact item within a narrowly defined consumption category as informative about the quality ladder (after accounting for potential savings from purchasing in bulk). This interpretation of prices as conveying quality information is similar to the logic of Jaravel (2019) who segments each product category into price deciles, interpreting these as representing a quality ladder.

The results of my event studies around non-employment have important implications for interpreting the expenditure drop during unemployment. Gruber (1997) was the first to empirically investigate the behavior of consumption over an unemployment spell. Recent work by Ganong and Noel (2019) and Landais and Spinnewijn (2021) has leveraged much richer data sets (bank account data and administrative records on earnings and wealth) to infer the expenditure drop around unemployment. The former find that non-durable expenditures at the onset of unemployment fall by around 8%, with declines of 1% for each additional month of unemployment and another 12% drop in expenditures at the exhaustion of unemployment insurance benefits. Relative to these magnitudes, the savings from shopping behavior are quite small.

The rest of this paper is organized as follows: Section 2 discusses the data and provides background on the categorization of expenditures in the Kilts Nielsen Consumer Panel. Section 3 lays out the economic logic of my shopping adjustment measures and 4 details how I map this to the data. In Section 5 I trace out the path of expenditures and shopping adjustments after retirement and show that relative to the drop in expenditures, shopping effects are small. Section 6 investigates to what extent shopping adjustments matter when comparing across vs. within-households. Section 7 discusses policy implications and avenues for future research.

2 Data and Variable Definitions

I use data from two panel surveys: the Kilts Nielsen Consumer Panel (henceforth "Nielsen Panel" or KNCP) and the the Health and Retirement Survey (HRS). The Nielsen Panel is a scanner data set that includes information on prices paid at the barcode-by-trip level and includes information on exact consumption bundles at grocery stores broadly defined. The HRS's Consumption and Activities Mail Survey (CAMS) provides information on virtually all Consumer Expenditure Survey categories of consumption and also elicits time use in a similar way to the American Time Use Survey.

2.1 Kilts Nielsen Consumer Panel

The main data set for this paper is the KNCP; a data set that covers an annual panel of households from 2004 to 2019. In total, the data set includes 194,551 households and

917,962 household-year observations. The KNCP includes date-exact information about purchases at grocery stores, supermarkets, discount stores, superstores and similar store categories. An annual set of about 60,000 households records purchases by logging each item, providing the exact UPC or providing additional information about data on goods like raw produce that is being sold by weight.[1] Purchases are recorded at the trip level, with information about the price paid, whether an item was recorded as being on sale (associated with a "deal"), whether a coupon was used, the exact number of items purchased and the overall expenditure for the shopping trip. Households are provided with financial incentives for their participation in the KNCP sample, and they may drop out of the sample at the end of a panel year or may continue from one year to the next. A panel year stretches from the last days of December of one year to mid-to-late December of the following year, which implies that a panel year aligns very closely, but not perfectly with the calendar year.

Once a year, in the fourth quarter preceding the data collection of data for a panel year, households are asked a variety of demographic questions. This includes household income, household size, whether a male and female household head are present, questions about household members' ages, education, occupation, a variety of information about living conditions (e.g., the kind of residence a household lives in, the availability of internet and TV service, the presence of a variety of household appliances). The exact date of the collection of this demographic information is not provided to researchers, so I will interpret all of these demographic variables as representing the data for the fourth quarter of the year these data were collected. Households may exit or enter the panel in any given year. About 80% of participants remain in the sample from year to year, and many households remain for the sample for substantial amounts of time. The average number of years in the panel for 4.7 years, and the average number of continuous years in the sample is slightly more than 4 years.

2.2 Product-Level Information

The main benefit of the KNCP is the level of granularity for purchases. Purchases and prices are recorded at the trip by UPC (barcode) level.[2]. While an increasing number of shoppers is asked to record all purchases (including items without an associated UPC code), throughout I will restrict attention to purchases with an associated UPC. There are

[1] Throughout, I will restrict attention to purchases associated with a UPC code.

[2] Sometimes, the same UPC will correspond to a different product in different years. Most notably, this is true for changes in the size or weight of a product. In these cases, Nielsen constructs a "UPC version" variable that assures that a UPC captures identical items. While I construct all my measures at the UPC-by-UPC version level, I will simply refer to this as "the UPC level" for expositional purposes.

about 3.5 million unique UPCs in the KNCP, which are mapped to 1,298 product modules in 110 product groups and 9 departments. Product modules are the lowest level of aggregation and correspond to very fine categories consumption. For example, frozen orange juice, fresh orange juice, fresh apple juice, and sugar-sweetened fruit beverages each are different product modules. Figure A.1 presents the product hierarchy in Nielsen, going from departments to product groups to product modules to individual UPCs. For every barcode, Nielsen provides information about the size or weight of a product, and the associated unit of measurement (e.g., ounces, milliliters, square feet, or counts). In addition, Nielsen records whether a UPC corresponds to a multi- or single-pack. This allows me to construct exact quantities for every UPC in the data. In turn, I can construct per-unit prices for each UPC to make prices comparable between products of different sizes. For example, one individual can of Coca-Cola, one two liter bottle of Coca-Cola and one 24-pack of cans of Coca-Cola will all be associated with different UPC codes. Using the quantity information provided by Nielsen, I can compute exact per-ounce of soda prices for each of these. Figure A.3 provides an example.

2.3 Baseline Sample

For the baseline sample, I only include households who are observable between the ages of 25 and 74 for at least five years so that I can estimate life-cycle profiles for the age range that has received the most attention in the existing literature.[3] Because I do not know who the "primary shopper" in a given household is, I define the age of a household as the average age of the household heads. I further restrict the sample to households for which I observe shopping trips in at least 11 months out of the year and real spending of at least $250 per year (in 2012 dollars as deflated by the CPI for food at home). This leaves me with 179,703 households and 814,938 household-year observations, where 121,553 households are observed for multiple years. In Column 1 of Table 1, I present summary statistics for this sample (weighted using KNCP projection factors). We can see that the weighted KNCP matches the US population quite well on most observables: the sample is broadly representative of the US as a whole in terms of income, household size, and other demographics like race or Hispanic origin (conditional on age).[4] In addition, about nine percent of households experience a retirement while in the panel, with another ten percent of households undergoing non-employment spells while in the Nielsen Consumer

[3]For example, Gourinchas and Parker (2002) construct life-cycle profiles of income and for ages 30 to 65. Aguiar and Hurst (2007) report information on prices paid for 5-year age bins ranging from 25 to 74.

[4]An important caveat is that the unweighted KNCP skews older and substantially over-represents female heads of household.

Panel.

2.4 Retirement Event Study Sample

As noted above, all demographic information is collected in the fourth quarter *preceding* the panel year. That is, if a household is in the Nielsen Consumer Panel for panel year 2004, then the associated demographic information will have been collected during the fourth quarter of 2003. Therefore, employment flows can only observed if the change in employment status covers the fourth quarter of any given year.[5] A second problem is that Nielsen does not actually record information about different kinds of non-employment, but rather reports a single non-employment category that combines unemployment, retirement, and voluntarily staying at home.

To identify retirements, I focus on household heads who have been employed in year $t-1$, who are between ages 60 and 70 in year t, and who are still not employed in year $t+1$.[6] Given these restrictions, all households in my retirement sample will have to be in the data for at least three consecutive years. I impose that household heads cannot "unretire": If I observe multiple retirements for a male or female household head, I only keep the first retirement. For non-concurrent retirement spells (that is, both heads of household retiring at different points in time), I treat the first retirement as the "treatment". Of these households, I then restrict the sample to household-year observations in which the household records at least one shopping trip for each month of the year. This leaves me with a sample of 10,007 households ever entering retirement. Control households are households with at least one working household head between the ages 55 and 70 for which I do not observe either a transition from employment to non-employment. We can think of these household as not yet retired or having at most one household head who retired before entering the Kilts Nielsen Consumer Panel.

In Column 2 of Table 1, I present summary statistics for the retirement event study sample. A household experiencing at least one retires undergoes 1.02 retirements on average and remains in the sample for 10.4 years. Otherwise, the retirement sample looks broadly similar to the baseline sample with most differences stemming from the fact that households in the retirement sample are substantially older. In Figure A.2, I present the

[5] A less consequential problem is that the exact date or even month during which the demographic information was collected is unknown.

[6] The logic for the age cutoffs is as follows. At age 59.5, workers are old enough to make penalty-free withdrawals from tax-advantaged retirement accounts (and often retire with pension benefits). From age 70 onward, there is no benefit to delaying claiming Social Security anymore. Therefore, I consider any age outside of this range as "unusual" or potentially driven by labor market shocks that aren't really about the life-cycle. Similarly, I require two years of consecutive non-employment to make sure my measure of retirements does not reflect unemployment spells.

distribution of retirement ages according to my assignment of retirements. Similar to cases with information on actual retirements, retirements spike at age 62—the earliest age at which workers become eligible for Social Security benefits. However, by construction my measure does not pick up the asymmetry of retirements present in data sets that explicitly elicit the age at retirement (the vast majority of retirements occur up until age 65 with relatively few retirements thereafter).

2.5 The Health and Retirement Survey

In addition to the Kilts Nielsen Consumer Panel, I will also leverage the Health and Retirement Survey and its Consumption and Activities Mail Survey module. The HRS is a biannual longitudinal panel of a nationally representative sample of households with heads ages 50 and older. I combine information on household demographics (household composition, age, race, ethnicity), household income, household assets, and retirement status from the main HRS sample with additional information from the Consumption and Activities Mail Survey (henceforth "CAMS"). CAMS households also provide information on very detailed categories of consumption (CAMS covers virtually all CEX categories of consumption) as well as time use similar to the American Time Use Survey. Given the rich information, this allows me to construct a bi-annual panel of households that includes information on assets, earnings, retirement status, consumption, and time use for the same household. My sample consists of all households in CAMS. In Table 2, I present summary statistics for not retired workers, retired workers, and the full sample. We can see that the sample is roughly comparable to the Nielsen Panel although CAMS households have somewhat lower incomes, fewer years of education, and are more likely to be a minority. The outcomes of interest in the HRS sample will be total expenditures, total non-durable expenditures, and time use on market and non-market activities.

3 Theoretical Framework

To motivate the empirical approach of this paper, I will present a simple model of shopping behavior that clarifies the empirical objects defined in the next section. Consider a household i that faces the following optimization problem

$$\max_{\boldsymbol{q},\boldsymbol{s},\boldsymbol{a}} u\big(c(\boldsymbol{q})\big) - h(\boldsymbol{s}) \text{ s.t. } e(\boldsymbol{q},\boldsymbol{s}) = \bar{y}, \tag{1}$$

where $\boldsymbol{q}$ denotes the households' consumption bundle, $c(\cdot)$ is an increasing concave function, $\boldsymbol{s}$ is shopping behavior with the disutility cost of lowering per-unit prices with $h(\boldsymbol{s})$ being an increasing convex function. In other settings $h(s)$ is often taken to be the opportunity cost of time, but it can principally be more general and capture storage costs or dis-utility from shopping at less nice stores. Lowering per-unit prices can be either achieved through paying less for a given item or by purchasing in bulk. e denotes expenditures and $\bar{y}$ is some exogenous budget constraint.[7] Expenditures are a function of the chosen bundle $\boldsymbol{q}$ and household-level prices (which depend on shopping behavior):

$$e_i = \sum_m \left(\sum_{k \in m} p_{i,k}(s_i) q_{i,k} \right) \tag{2}$$

where k denotes varieties of a given consumption category m, $p_{i,k}(s_i)$ is a household-specific price (where higher s_i implies weakly lower prices), and $q_{i,k}$ is household i's quantity of item k. Note that households do not derive utility from paying lower or higher prices outside of the relaxation of the households' budget constraint. However, within a consumption category m, households prefer higher-quality items which have higher prices *on average*. Therefore, $\bar{p}_k$, the average price of item k, is informative about the quality of item k relative to all other items $l \neq k \in m$. The utility from consumption of market goods is given by:

$$u(c_i) = u\big(c(\bar{p}_1 \cdot q_{i,1}, \ldots, \bar{p}_K \cdot q_{i,K})\big) \tag{3}$$

Therefore, we can define a measure of "consumption expenditure" that captures only the utility-relevant aspect of expenditures (leaving aside the dis-utility cost of finding the best deals or buying larger bulk quantities).

$$c_i^* = \sum_m \sum_{k \in m} \bar{p}_k q_{i,k}^*, \tag{4}$$

where $q_{i,k}$ are components of the consumption bundle $\boldsymbol{q}_i^*$ which is defined as

$$\boldsymbol{q}_i^* = \arg\max_{\boldsymbol{q}_i} u\big(c(\boldsymbol{q}_i)\big)$$

[7]This budget constraint could refer to rules of thumb such as consuming a constant fraction of income every period, but it may also differ across states of the world (e.g., retired vs. working, unemployed vs. employed).

Note that equipped with measure of consumption expenditure we can re-write 2 to get

$$
\begin{aligned}
e_i &= \sum_m \sum_{k \in m} \left(p_{i,k}(s_i) - \bar{p}_k\right) \cdot q_{i,t} + \sum_m \sum_{k \in m} \bar{p}_k \cdot q_{i,k} \\
&= c_i^* + \sum_m \left(\sum_{k \in m} \left(p_{i,k}(s_i) - \bar{p}_k\right) \cdot q_{i,k} \right)
\end{aligned} \tag{5}
$$

where $\sum_m \left(\sum_{k \in m} \left(p_{i,k}(s_i) - \bar{p}_k\right) \cdot q_{i,k} \right)$ is a wedge between consumption-relevant expenditure and expenditure. Household i may exert higher shopping effort so that it faces lower prices for each item k. In that case $p_{i,k}(s_i)$ will be less than zero and measured expenditure will understate the true amount of consumption expenditure for household i. Similarly, household i may allocate a large share of the budget to bulk purchases, thereby lowering overall costs. Here, too, $p_{i,k}(s_i)$ will be small, and measured expenditure for household i will understate consumption relevant expenditure for household i. The reverse holds for low levels of shopping for low prices or small bulk allocations.

Overall, this structure implies that a maximizing household will set $\boldsymbol{q}^*$ and $\boldsymbol{s}^*$ so that the marginal gain from relaxing the budget constraint will equal the dis-utility costs of increasing shopping effort and higher bulk allocations. For example, a household entering retirement may have lower opportunity cost of time so that $h(\boldsymbol{s})$ is lower for any level of s. This would result in that household exerting more shopping effort and paying lower prices. As children move out, a household may also have more available space and hence lower storage costs, again implying lower $h(\boldsymbol{s})$ for any $\boldsymbol{s}$. In that case, the household will buy more in bulk.

Additionally, when picking bundle $\boldsymbol{q}^*$, the household will pick quantities such that the marginal utility of improving the quality of the consumption bundle is equal to the marginal utility from increasing quantities.[8] This is particularly helpful when thinking about consumption of goods captured by the Nielsen Consumer Panel. Much of the covered items represent necessities so that we may not expect to see much adjustment in terms of total quantities. Movements in the quality of the consumption bundle are therefore informative about overall consumption adjustments (including in cases where we would expect a larger fraction of the adjustment to be accounted for by changes in *quantities*).

[8]Note that this need hold for allocating varieties k within a category of consumption m. More likely, the household will be indifferent between spending the marginal dollar on increasing the quantity of some category m and improving the quantity of some other category $n \neq m$.

4 Household-Level Prices and the Quality of Consumption

As outlined in the previous section, the main interest of this paper is to decompose expenditures into consumption expenditures c_i^* (themselves composed of quantity and quality effects) and shopping behavior that keeps the quantity and quality of the consumption bundle unchanged (but may lower nominal expenditures). In Section A, I show that we can express log expenditures as:

$$\ln e_i = \text{Quantity Effects}_i + \text{Price Effects}_i + \text{Bulk Effects}_i + \text{Quality Effects}_i \qquad (6)$$

Going back to the framework of the previous section, only Quantity Effects$_i$ and Quality Effects$_i$ are utility-relevant while Price Effects$_i$ and Bulk Effects$_i$ are sources of a wedge between observed expenditure and utility-relevant expenditure. Below, I will discuss how I measure each of these components in the scanner data.

Price Effects: Paying Lower Prices for Identical Goods

My way of measuring shopping effort builds upon the logic laid out by Aguiar and Hurst (2007): With UPC-level information, I can observe whether households pay lower prices for *identical* goods. I will construct cross-sectional price indices of consumption by comparing the realized prices a household actually paid to the identical bundle (at the UPC level) at average prices. Therefore, shopping effort for household i at time t is given by:

$$\text{Price Effects}_{i,t} = \underbrace{\ln\left(\sum_l P_{i,t,k} \cdot Q_{i,t,k}\right)}_{\text{Actual Cost of Bundle}} - \underbrace{\ln\left(\sum_l \frac{\sum_j P_{j,t,k} Q_{j,t,k}}{\sum_j Q_{j,t,k}} \cdot Q_{i,t,k}\right)}_{\text{Cost of Actual Bundle at Mean Prices}},$$

where $P_{i,t,k}$ denotes the price household i paid for UPC k at time t. In the data, purchases are recorded at the shopping trip level, which I aggregate to monthly, quarterly, or annual frequency, so t will correspond to monthly, quarterly, or annual date, with $P_{i,t,l}$ defined as the average price a household paid at that frequency. $Q_{i,t,k}$ is the quantity of UPC k household i purchased in period t. The second term computes the quantity-weighted average price for each UPC k in period t times the actual quantity of UPC k household i purchased in period t. I then normalize this variable to be centered at 0 every period by subtracting by its period-mean:

$$\text{Price Effects}^{*}_{i,t} = \text{Price Effects}_{i,t} - \frac{1}{N}\sum_{j}^{N}\text{Price Effects}_{j,t} \tag{7}$$

Intuitively, this measure of adjustment *only* captures prices paid relative to the average. Looking at Figure A.4 exerting shopping effort implies paying the pay prices for either store-brand conventional eggs or name-brand pasture raised eggs, not purchasing store-brand conventional eggs instead of name-brand pasture-raised eggs. However, since the price index is defined cross-sectionally, a household need not purchase the same bundle every period. For example, a household could pay exactly the average price for store-brand conventional eggs one period, and then pay exactly the average price for name-brand pasture-raised eggs in the next period. In both cases, the corresponding effort margin would be equal to 0. In Column 1 of Table 3, I present moments of the distribution of the effort margin. The standard deviation is around 8.1% while the inter-quartile range is approximately 7.9%. This suggests that there is substantial room to exert shopping effort in order to lower prices.

Bulk Effects: Purchasing Larger Quantities

Another potential way to reduce the per-unit cost of their purchases is for households to purchase in bulk. Note that unlike paying less for a given UPC, this channel of lowering prices requires some storage costs. As such, it is not entirely clear if we should think of bulk-purchases as a pure reduction of per-unit costs or as a costly way of lowering the price of the consumption bundle. When constructing my measure of bulk savings, I proceed as follows: Following Griffith et al. (2009), I break out each product module into five quintiles of the size distribution. I then define UPCs in the top two quintiles (or the Top 40% of the within-product-module size distribution) as bulk items.[9] I then define bulk savings as follows:

$$\textit{Bulk Effects}_{i,t} = \underbrace{\ln\left(\sum_{m\times b}\left(\frac{\sum_{k\in m\times b}\sum_j P_{j,t,k}Q_{j,t,k}}{\sum_{k\in m\times b}\sum_j Q_{j,t,k}}\cdot\sum_{k\in m\times b}Q_{i,t,k}\right)\right)}_{\text{Cost of Bundle at Actual Bulk Share}}$$

[9]Note that I use *total quantities* to define item size. For example, a twelve-pack of cans of Coca-Cola will be defined as equaling 144 fluid ounces (12 times 12 fluid ounces), roughly similar to two-liter bottles of Coca-Cola (135.2 fluid ounces).

$$- \underbrace{\ln \left(\sum_m \left(\frac{\sum_k \sum_j P_{j,t,k} Q_{j,t,k}}{\sum_k \sum_j Q_{j,t,k}} \cdot \sum_k Q_{i,t,k} \right) \right)}_{\text{Cost of Bundle at Average Bulk Share}},$$

where i and j denote households, m indexes product modules, b indexes bulk and non-bulk, t indexes time and k denotes UPC-level items. Intuitively, bulk savings are the difference in the cost of the purchased bundle at actual bulk shares (for each product module separately) and the cost of the bundle had the household purchased it at "average" bulk shares, both evaluated at average per-unit prices for a given product-module-by-bulk-level combination. In that sense, the bulk savings measure abstracts away from actual prices—all savings arise solely from allocation a higher budget share to bulk items (assuming bulk items are cheaper). I normalize this variable to be centered at zero every period by subtracting its period-mean:

$$\text{Bulk Effects}^*_{i,t} = \text{Bulk Effects}_{i,t} - \frac{1}{N} \sum_j^N \text{Bulk Effects}_{j,t} \tag{8}$$

In Column 2 of Table 3, we can see that the potential savings from bulk purchases are quite meaningful: The standard deviation is 6.1% and the interquartile range is 7.4%. This is particularly true in the tails where moving by just 5 percentiles (i.e., from p95 to p90 or from p10 to p5) results in approximately 2.5% lower prices for a given bundle.

Going back to the framework in Section 3, it will be convenient to think of the bulk and effort margins as the overall wedge between observed expenditure and consumption expenditure. I also define this aggregate savings measure as follows:

$$\text{Expenditure Wedge}_{i,t} = \text{Price Effects}^*_{i,t} + \text{Bulk Effects}^*_{i,t} \tag{9}$$

Since the expenditure wedge combines the two measures of adjustment that keep the quality of the bundle fixed, looking at the distribution of the expenditure wedge gives us a sense of the maximum possible savings through paying lower prices for a given UPC and purchasing a higher fraction of bulk items. The standard deviation of the expenditure wedge is 9.9% and the interquartile range is 11.9%.

Quality Effects: Purchasing Goods with Lower Average Prices

My definition of quality adjustments leverages the richness of the Nielsen Panel. Within a given product module, households can move down the quality ladder and purchase

UPCs that have lower prices *on average*. This measure of quality abstracts away from the realized prices any given household pays and compares the average per-unit price of the items the household actually purchased to the average per-unit price of that product-module-by-bulk level. To make this explicit, quality adjustments are given by

$$\text{Quality Effects}_{i,t} = \underbrace{\ln\left(\sum_m \sum_k \frac{\sum_j P_{j,t,k} Q_{j,t,k}}{\sum_j Q_{j,t,k}} \cdot Q_{i,t,k}\right)}_{\text{Cost of Actual Bundle at Mean Prices}} \\ - \underbrace{\ln\left(\sum_{m\times b}\left(\frac{\sum_{k\in m\times b}\sum_j P_{j,t,k} Q_{j,t,k}}{\sum_{k\in m\times b}\sum_j Q_{j,t,k}} \cdot \sum_{k\in m\times b} Q_{i,t,k}\right)\right)}_{\text{Cost of Average Bundle at Mean Prices}_{i,t}},$$

where m denotes product modules, b denotes bulk vs. non-bulk, and k indexes UPCs within a product m times bulk level b. As above j indexes households other than i, t indexes time at monthly, quarterly, or annual frequency, and P and Q denote prices and quantities, respectively. Just like with shopping effort, I normalize the quality measure so that it is centered at 0 in every period:

$$\text{Quality Effects}^*_{i,t} = \text{Quality Effects}_{i,t} - \frac{1}{N}\sum_j^N \text{Quality Effects}_{j,t} \tag{10}$$

To illustrate the logic of this measure of quality, consider the example in Figure A.6. A every period, the household can move along the quality ladder by purchasing store-brand eggs, name-brand cage-free eggs, store-brand organic eggs, or name-brand pasture-raised eggs. The quality effect metric will compare the cost of eggs at the average price of the eggs actually purchased to the cost of the same number of eggs at the average price of eggs in the same period. As I show in Table 3, the quality effects are more variable than even the combined price and bulk effects: The standard deviation of the quality metric is 13.3% and its interquartile range is 16.4%.

In principle, a household could pay very low prices for items that tend to be very expensive on average (e.g., through buying in bulk or waiting for deals, by going to different stores, or by using coupons). Similarly, a household may pay very high prices for eggs that tend to be cheap on average, for example by purchasing eggs at a corner store or by buying eggs that are frequently discounted at full price. In practice, paying higher prices (a high effort margin) and purchasing higher quality items are positively correlated (a correlation of 0.226 at annual frequency). Similarly, purchasing higher quality bundles is correlated with purchasing smaller quantities (0.121 at annual frequency) while buying in bulk and lowering UPC-level prices are not correlated at all.

Validating the Quality Measure

Given the distinction between prices and quality, it is important to validate that my quality metric really does represent differences in the quality of the consumption bundle. For example, one might be concerned that my measure of quality effects overstates differences in product quality and really picks up differences in *prices*.[10] In this section, I will provide evidence that my quality measures really do pick up *quality* differences, not just price effects.

A natural question to ask is whether the my quality metric varies by permanent income. For this purpose, I assign each household an earnings rank based on their average earnings within a given cohort.[11] As we can see in Panel (a) of Figure 1, bundle quality is monotonically increasing in lifetime income rank, despite the fact that the quality metric is constructed without levering any information on income (or even on the total amount of expenditure). In addition, the slope of the quality measure that is constructed locally is almost identical to the slope of a national comparison. This suggests that my quality measure picks up true variation in purchasing behavior rather than a positive correlation between local incomes and prices.

Prices paid, on the other hand, are not very strongly related to lifetime income. The slope of the prices and lifetime income profile follows a rough u-shape: Prices paid relative to average are *falling* from the 1st to the 20th percentile, then they are relatively flat from the 15th to 40th percentile, slowly increasing from the 40th to 80th percentile, and then the are rapidly increasing for the highest lifetime income percentiles. For local prices, this pattern in generally true as well. However, the increase in prices for the highest income percentiles is much starker for national than local prices. This suggests that national price comparisons do pick up some of the covariance between local incomes and local prices. Regardless of this fact, the range of prices paid is much more narrow than that of quality and lifetime income ranks explain more than an order of magnitude more of the variation in quality than prices paid.[12]

[10] It should be noted that the opposite might be true as well. Using mean UPC-prices may treat quality differences as differences in prices. For example, the same (in a UPC-code sense) gallon of milk could be purchased at Whole Foods or a large discount store. To the extent that the shopping experience at Whole Foods is more pleasant than at the discount store, this may reflect *quality* and not price differences. A similar argument can be applied to commute time.

[11] This way, my earnings ranks are only about household earnings ranks within their age group instead of picking up the age profile of earnings.

[12] Regressing the quality and price measures on lifetime income ranks (either linearly or on 100 income

As a second test show the relationship between my quality and price measures and county-level unemployment rates. Controlling for household characteristics, time and market fixed effects local unemployment rates are highly correlated with the quality of consumption. Going from a county-level unemployment rate of 5 percent to a rate of 10 percent, the quality of the average consumption bundle falls by 5 percentage points. Effective prices, on the other hand, are not strongly correlated with local economic conditions. If anything, it seems like effective prices are *increasing* in the local unemployment rate for unemployment rates above 5 percent. One potential explanation for this is that households in areas with very high unemployment have less access to cars that make cheaper stores more easily accessible.

Finally, we can check to what extent the measures of consumption are correlated with household level measures of socio-economic status. In Table 4, I perform six such tests. In the first column, I regress the quality measure on log household income and a set of household level controls. We can see that a one percent increase in current income increases the quality of consumption by 0.06%. To test to what extent current or lifetime income drives this effect, I regress quality on lifetime income rank and current income in column (2). We can see that the coefficient on current income falls by about two-thirds while a one percentile increase in permanent income increases the quality metric by 0.14 percent. From this, we can infer that the quality of the consumption bundle is driven both by permanent and current income. To test whether wealth predicts the quality of consumption, I regress the quality measure on the ZIP code-level house price from Zillow for the subsample of households living in a single family home, a condo, or a co-op.[13] We can see that higher house prices go along with substantially higher quality of consumption, even conditional on income. In the fourth column, I leverage a matched data set between the HomeScan data and the Survey of Consumer Finances (details on the matching procedure can be found in Appendix Section B.1). Based on their SCF matches, I group households into within-cohort net-worth ventiles and regress the quality of consumption on the imputed rank in the wealth distribution. Higher imputed net worth ranks result in substantially higher consumption quality according to each measure. Conditional on income, going from the 5th to the 15th ventile (so the 25th to the 75th percentile) of the within-cohort wealth distribution results in consuming module bundles of 1.5% higher quality. In the fifth column, I regress the quality measure on five educational attainment

rank indicators) yields an R^2 of 0.0629 and 0.0643 for the quality metric but only an R^2 of 0.0021 and 0.0054 for the effort metric.

[13] Unfortunately, the Nielsen Panel data does not elicit home-ownership directly but only asks household whether they live in a single family home, a two-party home, a multi-family home, or a mobile home. Then the survey separately elicits whether households live in a condo or co-op.

indicators: less than high-school, a high-school degree, some college, a college degree, or a post-graduate degree. The quality of consumption is monotonically increasing in educational attainment, even conditional on current income. This is consistent with the fact that more educated households are likely to have higher lifetime incomes (even conditional on current income). In the final column, I include all of these variables at the same time. All coefficients are statistically significant and economically meaningful. Current income and current wealth (as measured by house prices and imputed wealth based on the SCF) both predict the quality of consumption; and so do lifetime income ranks and household-level educational attainment. Given that household-level education is highly predictive of the quality of consumption even conditional on current income, lifetime income, and imputed wealth suggests that my quality metric also captures the *preferences* of more educated households to some extent. For example more highly educated households may prefer organic pasture-raised eggs to battery-cage eggs at any level of income. Throughout, I will include a household fixed effect, so none of my results are driven by this potential for preference heterogeneity.

5 Decomposing the Post-Retirement Expenditure Drop

In my first set of results, I will focus on dynamic adjustments around retirement. These allow me to follow the same households as it enters retirement and trace out their expenditure profile, cost savings arising from paying lower prices and buying bulk as well as the quality of consumption. My general empirical strategy for the event studies is the following framework:

$$\text{Outcome}_{i,t} = \theta_i + \sum_{\tau=-A}^{B} \mathbb{1}\{T_{i,t} = \tau\} \cdot \delta_\tau + \boldsymbol{X}_{i,t}' \gamma \epsilon_{i,t}, \tag{11}$$

where θ_i = is a household fixed effect, δ_τ = are leads and lags for last known date of employment (before retirement or unemployment), and $\boldsymbol{X}$ is a vector of household composition controls (household size, the relationship of the adult heads of household, presence of children). The coefficient of interest are the δ_τ which tells us how the outcome of interest evolves relative to the last year before retirement.[14]

[14]Note that I am not including time fixed effects in my baseline results. The retirement-consumption puzzle is about falling expenditures for a given household, not falling expenditures relative to a growth trend. However, I also estimate a two-way fixed effect version of my results in the appendix. Recent work (Borusyak et al. 2022, Callaway and Sant'Anna 2021, de Chaisemartin and D'Haultfœuille 2020, Sun and Abraham 2021) has shown that estimating models like the one above with ordinary least squares will not

5.1 Income, Expenditures, and Shopping Behavior

Throughout, I will present dynamic event study coefficients where coefficients are plotted relative to the last year in which the household member was still employed so that we know that the retirement occurred at some point in period 0.[15] As we can see in Figure 2, household income is flat or trending down before retirement and then falls abruptly. An important note is that the income drop is about twice as big in the HRS as it is in the Nielsen Panel. Potential explanations for this are the topcoding of high incomes in the HomeScan data, the fact that higher-educated households are over-represented in the Nielsen Panel, or measurement error in the income measure discussed in Section 2. In Figure 3, we can see that HomeScan-covered expenditures and the broader set of HRS-covered expenditures follow broadly similar trajectories, particularly after retirement. Six years into retirement, scanner-covered expenditure has fallen by about 8 percent while total non-durable expenditure in the HRS has fallen by around 15 percent. In both cases, this suggest some smoothing by households as incomes fall by substantially more than expenditures. Nonetheless, the marginal propensity out of the post-retirement income drop is large and the declines expenditure are economically meaningful.

Turning the question how households adjust their shopping behavior in response to retirement, Panels (a) and (b) of Figure 4 show event study results for savings arising from paying lower UPC-level prices and purchasing larger quantities (i.e., bulk savings). We see that there is very little evidence of pre-trends for prices paid relative to the average while bulk savings fall smoothly through retirement. Four years into retirement (at the trough of prices paid and one year after the trough for bulk savings), households pay around 0.2 percent lower prices for a fixed UPC than they did just prior to retirement while they save an additional 0.2 percent on per-unit prices by changing their bulk allocation. The quality of the consumption bundle, on the other hand, changes significantly. As we can see in Panel (c), households rapidly substitute towards cheaper, lower-quality items within narrowly defined categories of consumption. Four years into retirement, quality of the purchased bundle relative to the average bundle has fallen by about 1.2 percentage points.

These results have stark implications. In Panel (a) of Figure 5, I decompose the change in the household's deviation from the cost of the "average" bundle (that is, a bundle con-

generally yield unbiased results, particularly in the presence of treatment effect heterogeneity. Therefore, I will actually estimate this model using the estimator proposed by de Chaisemartin and D'Haultfœuille (2020) in my baseline specification.

[15] Since the HRS is a bi-annual survey, this means that we can only group the leads and lags into two year bins. Therefore period -4 refers to four to three years prior to retirement, $t = -2$ refers to two to one years prior to retirement, $t = 0$ refers to zero to one years since retirement, and so on.

sisting of the same consumption categories, but at average prices, average bulk shares, and average quality). We can see that bundle quality, rather than price and bulk effects, is the main margin of shopping adjustment. Quality adjustments are three to four times as large as the combined adjustment from paying lower prices for a given good or increasing bulk purchases. This suggests that the main shopping adjustment is one that is costly in utility terms rather than one that yields similar levels of utility at lower expenditures through lower prices. In Panel (b), I present the path of expenditures, the path of *price and bulk-corrected* expenditures, and the path of quantities only (so declines in expenditure that arise from purchasing fewer items or substitutions across product modules). Consumption expenditures fall by almost as much as uncorrected expenditures. There also is very evidence of households "learning" how to save money: the wedge between corrected and uncorrected expenditures *falls* later into retirement and has fully disappeared six years into retirement. Even accounting for downward adjustments in item quality, quantities fall by almost 7%. Taken together, these two pieces of evidence imply that the post-retirement expenditure drop is a real drop in consumption expenditures, with utility-constant shopping effects explaining *at most* 5 to 10 percent of the expenditure drop.

5.2 Home Production and Expenditure Timing

An important caveat is that the estimates in Figure 5 still only capture *market* consumption. To the extent that households cut durable goods and work-related expenses and adjust their consumption bundles towards more home production, *expenditures* will overstate the consumption drop. In order to test these channels, I perform four separate tests: First, I make use of the detailed nature of the Scanner data and group goods into three categories: Goods that cannot be substituted with home production (e.g., shampoo, trash bags), goods that are substitutes for consumption away from home but are not inputs for home production (e.g., ready-to-eat foods), and home production inputs (e.g., unprepared produce and meat, flour, fresh eggs). I then estimate the expenditure drop for each category separately. As we can see in Panel (a), spending on non-substitutable goods falls by more than spending on home-consumption goods and spending on home production inputs do. However, spending on home consumption and home production inputs also falls. Given that expenditure on these categories also falls, at-home consumption and home production cannot offset declines in consumption elsewhere. In Panel (b), I conduct a similar exercise and leverage the fine-grained information in the CAMS module of the HRS to construct expenditure variables for home production goods (food at home, cleaning products, gardening products) and their immediate market good substitutes (food away from home, cleaning services, gardening services). Importantly, CAMS also includes information on time spent in each of these activities (time spent cleaning, doing gardening work, preparing meals and cleaning up and time spent shopping or running errands). In Panel (b) of Figure 6, we can see that expenditures on home production inputs fall almost as much as expenditures on the equivalent market goods. Therefore, it seems unlikely that home production could offset the decline in spending on market production. As a complement to this, we can consider households' time use: As we can see in Panel (c), the total time engaged in market work per week decreases by about 22 hours at retirement. However, the time spent preparing meals, shopping, cleaning, and gardening increases by only 4.5 hours for the first two years after retirement and then slowly declines thereafter.

Last, I investigate to what extent the large decline in total expenditures in the HRS could be the result of expenditure timing (for example, households purchasing new durables right before retirement, thereby elevating pre-retirement expenditures). To test this, I make use of the RAND HRS CAMS Data File that constructs a measure of *consumption* that accounts for principal repayment of mortgages and the fact that the consumption of durables occurs over time rather than all at once.[16] As we can see in Panel (d) of Figure 6, consumption spending does fall by less than total spending, but the declines in consumption expenditure are still substantial and there is no evidence of a spike in expenditures *before* retirement. The declines in total spending and are substantially larger than declines in non-durable spending alone (with consumption spending falling in between the two).

5.3 Mechanisms and Heterogeneity

An important question is how much the adjustments at retirement vary across households. Hurst (2008) argues that many studies of the post-retirement expenditure drop actually find *zero* adjustments at the median so that most of the drop is driven by a relatively small set of households who might be myopic and did not sufficiently plan for retirement. On the other hand, Bernheim et al. (2001) find that even households with relatively low declines in post-retirement income or high wealth cut their consumption after retirement. In their setting, households with low assets and larger post-retirement income drops do experience the largest declines in expenditures, but the expenditure drops are ubiquitous across the income and wealth distributions (the only group not seeing an expenditure decline is the set of high-post-retirement-income high-asset households).

[16] In particular, the RAND HRS CAMS Data File uses information from the Consumer Expenditure Survey to separate out interest expenses and principal payments on mortgages. For durables, it applies a per-period flow usage transformation following the approach of Hurd and Rohwedder (2006).

To tackle the question of the heterogeneity of the consumption drop, I perform several tests in the spirit of Bernheim et al. (2001). First, I group households into terciles according to their post-retirement income drop. In order to not pick up any one-time fluctuations, I define the income drop as the log difference in the average income over the three years preceding and immediately succeeding retirement. I then estimate the expenditure drop, the drop in consumption expenditure, as well as the expenditure wedge and quality adjustments for each tercile of the income drop. Similar to the finding of Bernheim et al. (2001), expenditures fall for every tercile of the income drop. Nonetheless, the expenditure drop for households with the smallest income drop is only about half as large as that for the two other terciles, suggesting an important role for current income. The wedge between expenditure and consumption relevant expenditure falls significantly only for households with the largest income drop, but even these declines are very small relative to the overall decline in expenditure. The quality of the bundle, too, declines the most for households with the largest income drop, with quality adjustments among the first tercile about three to four times the size of the adjustment for the top tercile.

As a second test, I match the Nielsen Consumer Panel to the Survey of Consumer Finances to impute household wealth (see Section B.1 for details) and group households by their pre-retirement imputed wealth. Here, the findings are also resembling those of Bernheim et al. (2001): Expenditures fall for each tercile of imputed wealth, but they fall the most for households with the lowest pre-retirement wealth. Savings arising from paying lower purchases and more bulk purchases are largest for the middle tercile with very little adjustment for either the top or the bottom tercile of pre-retirement wealth. Bundle quality, on the other hand, falls the least for the top tercile of wealth, with the drops for the bottom two terciles looking quite similar.

In results not reported here, I further investigate heterogeneity by educational attainment and pre-retirement income. Results are similar in the sense that it is households with lower incomes and lower education attainment who see the largest declines in expenditures, corrected expenditures, and bundle quality. This suggest that price and bulk effects at retirement are not only small on average, they are also small across many socioeconomic observables. On the other hand, households with lower current income, households with lower permanent income, and households with lower levels of wealth are all more affected by falling expenditures at retirement. These results are also borne out in the CAMS data which has joint information on assets, income, and expenditures. Total expenditure falls by much more for households with low wealth-to-income ratios or large post-retirement income drops. In terms of non-durables, point estimates are generally negative, but in most cases, I cannot reject no declines in expenditures for households with high wealth or a small post-retirement income drop (see Appendix Figure 9). From a social welfare perspective, these groups are likely to have relatively high marginal utilities, suggesting that the post-retirement expenditure drop also has important consequences for aggregate welfare.

Together, these results suggest that households really are differentially insured against the prospect of retirement. Some households see much smaller declines in current income that also go along with much smaller drops in expenditure. Wealthier households are also much better insured against large declines in current income after retirement, suggesting that differences in wealth accumulation are an important factor in explaining the heterogeneity in the magnitude of the post-retirement expenditure drop. These results are similar to the findings of Ganong and Noel (2019) and Ganong et al. (2020) in the context of unemployment and general fluctuations in labor earnings. In their setting, households with low liquidity are much more sensitive to unemployment or other labor-demand driven fluctuations in earnings. Considering more extreme cases, Ganong and Noel (2022) show that 70% of mortgage defaults are driven by adverse life events (shocks to current and future income) rather than negative equity or the interaction between negative equity and adverse life events. My results indicate that liquidity is an important driver of consumption behavior not just in light of these unanticipated shocks but even in light of anticipated shocks to current income. This points towards an important role for present bias or mental-accounting consumption behavior that puts weight not just on permanent income but *current* income as well.

5.4 Robustness

One potential concern with my estimates is that my main measures of shopping adjustment are all defined at the national, rather than local level. To investigate to what extent this has an effect on my results, I re-estimate the event studies for price effects, bulk effects, and quality adjustments making only local comparisons when constructing my shopping measures. For this, I leverage the fact that Nielsen divides its panel in to markets. These are 76 areas, which sometimes align with metro areas but need not be contiguous (e.g., rural counties surrounding a metro area might be defined as one market while the urban core could be defined as another). As we can see in Figure A.9, the price and bulk effects are *smaller* when using local comparisons while the quality effects are very similar to the effect estimated using national comparisons.

I further explore whether my event study results are robust to my choice of estimator. In order to do so, I re-estimate my event studies using the estimator proposed by de Chaisemartin and D'Haultfœuille (2020). As we can see in Figures A.11 and A.12, the income and expenditure drops are a bit smaller. However, expenditures still discretely drop after retirement. The estimated shopping effects are very similar to those of my baseline specification with price effects being a bit larger and bulk effects being a bit smaller (see Figure A.13). Overall bundle savings are almost identical to my baseline specification as are quality adjustments (see A.14). Taken together, shopping effects explain at most 10% of the post-retirement expenditure drop so the result that the post-retirement expenditure drop is not driven by shopping effects is robust to the choice of estimator.

6 When Do Shopping Effects Matter?

Given the results of the previous section, it is worth revisiting the question when shopping effects matter for household level prices and the quality of the consumption bundle. In order to do so, I report the effect of household head labor market attachment in the spirit of Kaplan and Menzio (2015). An important deviation from their estimates is that I break out labor market attachment for male and female household heads separately. In Panel A of Table 6, I report estimates for the set of households with two household heads and an "average" household age between 25 and 54. In Column (1), we can see that households with a *female* head not working full time are paying around 0.9% lower prices while the male heads' labor market participation has no effect on household-level prices. Controlling for household income, the effect for female household heads gets a bit smaller, although households with female household heads working part time or not working at all are still paying around 0.75% lower prices. Including a household fixed effect, the coefficient on the female heads' labor market attachment gets cut in half for non-employment and falls by a factor of four for female heads working part time.

With respect to the quality of the consumption bundle, we can see that much of the variation explained by household heads' labor market attachment is the result of the effect on earnings. While households with fewer employed household heads consume much cheaper bundles, conditional on income, this effect disappears or even reverses. Including a household fixed effect, the effect of labor market attachment on the quality of the consumption bundle is still statistically significant, but of much smaller economic magnitude than for the specification without household fixed effects or income controls.

In Panel B, we can see that a similar story applies for household ages 55 to 74. Unlike for prime-age households, the male head's labor market attachment now is predictive of household-level prices. Households with a female head not working full time pay between 0.7% and 1% lower prices, while households with a male head not working full time pay between 0.4% to 0.5% lower prices. Controlling for household income, all coefficients shrink a bit, but most are still statistically significant. Including a household fixed effect, male heads' labor market participation is no longer predictive of household level prices. Households with female heads not working full time pay around 0.2% lower prices, a similar magnitude to the maximum effect of retirement on household level prices estimated in Section 5. For bundle quality, patterns are quite similar: Households with heads not working full time are purchasing significantly cheaper items, an effect almost entirely driven by the effect of work on household income. Including a household fixed effect, not working full time reduces the quality of the consumption bundle by 0.6 to 0.75% with no effect of part time work for male household heads.

Together, these results suggest that the division of labor and household-level labor supply choices play an important role in explaining household level prices. If household-level prices were only a measure of household's opportunity cost of time, we would expect male household heads' employment status to also be an important determinant of household-level prices. However, most coefficients on male heads' employment status are not significantly different from zero and reasonably precisely estimated. More generally, a lot of the dispersion of household level prices can be explained by household fixed effects rather than within-household across time variation.

The Case of Unemployment

Given the apparently large role of fixed household characteristics in terms of explaining household-level prices, it is instructive to investigate the adjustment around a shock different from retirement: unemployment. In some ways retirement and unemployment are quite similar: Both retired and unemployed households have more free time to engage in money-saving shopping activities or home production.[17]

However, there are also very important differences for these two shocks. Unemployment represents a (mostly) unanticipated shock to income. Retirement, on the other hand, is (mostly) anticipated. Therefore, ex-ante it seems likely that there are more precautionary shopping effects for households entering retirement rather than households entering an unemployment spell. In addition, unemployment spells decrease lifetime in-

[17] An important note is that households facing unemployment will also spend some of their time looking for a job. Work by Krueger and Mueller (2010) suggests that the time spent looking for a job would still leave ample time to exert shopping effort to lower household-level prices. Households facing adverse labor market shocks also have other adjustment margins. For example, Koustas (2018) finds that ride-share drivers a substantial fractions of lost earnings in primary jobs with ride-share earnings, suggesting that flexible labor supply adjustments may be a very important aspect of households' self-insurance behavior against adverse labor market shocks.

come (a result going back to Jacobson et al. 1993), while retirement should leave lifetime income unchanged. Therefore, we would expect larger adjustments for unemployment than retirement. A factor pushing in the opposite direction is that unemployment is (usually) temporary while retirement is an absorbing state. To the extent that households consumption and shopping behavior are driven by *current* income, this would result in larger adjustments for households entering retirement.

In Figure 10, I present event studies for households undergoing non-employment spells. Since demographics are only elicited once a year, I cannot directly observe when households become unemployed. Rather, I center all coefficients at the last quarter before the unemployment spell so that $t = -1$ corresponds to the last known quarter of employment and $t = 4$ corresponds to the quarter in which a household member was non-employed. To make sure that most of the non-employment I am picking up is unemployment (rather than people voluntarily withdrawing from the labor force), I further restrict attention to unemployment spells that end after no more than two years. Estimating the effects on a balanced window, we see that during these non-employment spells, the brunt of the shopping adjustment falls on quality changes. Between the last quarter of known employment and the first quarter of known non-employment, bundle quality falls by about 1.1%. Prices paid fall by around 0.3% over the same time horizon while savings from bulk allocations are unresponsive to the non-employment spell. A notable feature of the adjustment is that effects on household-level prices and bundle quality revert back over time, though only prices recover fully. One potential explanation for the persistence of the effect on bundle quality is that unemployment tends to go along with declines in life-time income.

Implications for Using Expenditure as a Proxy of Consumption

Together, my results suggest that an important determinant of household-level shopping behavior is a household or type fixed effect. Households with female heads who are not working full time pay substantially lower prices, but this appears to be largely a result of differences in the household production function between different households. Within households, household heads' employment status does not affect household-level prices much. In the case of male heads of household, their employment status rarely matters for household-level prices irrespective of the choice of estimation procedure. This either implies that men and women have substantially different opportunity costs of time or that shopping is part of a larger intra-household bargaining problem for which opportunity costs of time are just one of many considerations. Another implication of these patterns it that shopping effects matter a lot for cross-sectional comparisons, but much less for within-household-across-time comparisons. That means that analyses relying on expenditures as a proxy for consumption are likely to be good approximations of true consumption adjustments as long as the underlying variation is within household.

7 Conclusion

Is the post-retirement expenditure drop a true drop in consumption? Decomposing expenditure into shopping behavior that keeps the consumption basket constant and quality and quantity adjustments that are costly in utility terms, I find that at least 90% of the expenditure decline after retirement represent true declines in quantities or quality adjustments. I further investigate whether substitution towards home production or expenditure timing can explain the expenditure drop. Expenditures declines are ubiquitous across all sub-components of consumption and even expenditures on home production inputs fall in retirement. I enrich these event study results by estimating the life-cycle profiles of household-level prices for identical items, bulk savings, and the quality of consumption. In stark contrast to prior literature, I find that household prices have a hump shape. Households do pay lower prices in old age than they did in their peak earnings years, but the difference between prices paid in households' 60s versus their 40s is only 0.3 to 0.8 percent. The difference between my estimates and that of prior work are primarily driven by the fact that my long panel allows me to trace out within-household life-cycle profiles instead of having to rely on cross-sectional comparisons. Taken together, these results suggest stark departures from the life-cycle permanent income hypothesis.

These patterns have important policy implications: The high sensitivity of households' consumption to current income implies that the payout scheme is a crucial policy lever when designing pension systems (irrespective of their funding mechanism). The main source of a constant stream of income for households in retirement is Social Security which only replaces 40% of pre-retirement earnings on average. One potential remedy for the large post-retirement expenditure drop is to increase replacement rates of Social Security, particularly at the lower end. More generally, my findings suggests that retirement schemes that guarantee constant cash flows may have substantial welfare benefits relative to systems without recurring payouts. This question of optimal payout schemes is particularly important as future retirees will be ever more likely to be drawing from defined-benefit plans. Given the failure of many households to annuitize their wealth, policies to increase annuity take-up or other ways to derive stable income flows from private retirement accounts are likely to have substantial welfare benefits. This is particularly important as the US transitions from a private retirement system mostly composed of defined benefit plans with guaranteed income flows to one dominated by defined contribution plans without any pre-set withdrawal strategies.

My results suggest multiple avenues for future research. On the empirical side, it would be interesting to use richer information on assets and income to explore in more detail which households are more insulated from the post-retirement expenditure drop. This work could then explore heterogeneity in the mechanism underlying the expenditure drop. For some households, it may be the result of low-wealth and a large decline in income, forcing the large adjustment. For wealthier households, important avenue for future research will be to disentangle to what extent the low rates of dis-saving are driven by large bequest motives, a failure to annuitize wealth holdings, or high sensitivity to current income even among the wealthy. On the theoretical side, the most common explanation for a high sensitivity of consumption to current income are high returns on illiquid assets or a combination of liquidity constraints and present focus. However, both classes of models are hard to reconcile with drops in expenditures in retirement even for relatively wealthy households. Models of consumption that can rationalize these behaviors would be an important contribution to our understanding of households' consumption behavior.

References

Aguiar, M. and Hurst, E. (2005). Consumption versus Expenditure. *Journal of Political Economy*, 113(5):919–948.

Aguiar, M. and Hurst, E. (2007). Life-Cycle Prices and Production. *American Economic Review*, 97(5):39.

Banks, J., Blundell, R., and Tanner, S. (1998). Is There a Retirement-Savings Puzzle? *The American Economic Review*, 88(4):769–788. Publisher: American Economic Association.

Bernheim, B. D. (1987). Dissaving after Retirement: Testing the Pure Life Cycle Hypothesis. In *NBER Chapters*, pages 237–280. National Bureau of Economic Research, Inc.

Bernheim, B. D., Skinner, J., and Weinberg, S. (2001). What Accounts for the Variation in Retirement Wealth among U.S. Households? *American Economic Review*, 91(4):832–857.

Borusyak, K., Jaravel, X., and Spiess, J. (2022). Revisiting Event Study Designs: Robust and Efficient Estimation. arXiv:2108.12419 [econ].

Callaway, B. and Sant'Anna, P. H. C. (2021). Difference-in-Differences with multiple time periods. *Journal of Econometrics*, 225(2):200–230.

Cattaneo, M. D., Crump, R. K., Farrell, M., and Feng, Y. (2021). On Binscatter. SSRN Scholarly Paper 3344739, Rochester, NY.

Coibion, O., Gorodnichenko, Y., and Hong, G. H. (2015). The Cyclicality of Sales, Regular and Effective Prices: Business Cycle and Policy Implications. *American Economic Review*, 105(3):993–1029.

de Chaisemartin, C. and D'Haultfœuille, X. (2020). Two-Way Fixed Effects Estimators with Heterogeneous Treatment Effects. *American Economic Review*, 110(9):2964–2996.

Ganong, P., Jones, D., Noel, P. J., Greig, F. E., Farrell, D., and Wheat, C. (2020). Wealth, Race, and Consumption Smoothing of Typical Income Shocks.

Ganong, P. and Noel, P. (2019). Consumer Spending during Unemployment: Positive and Normative Implications. *American Economic Review*, 109(7):2383–2424.

Ganong, P. and Noel, P. (2022). Why do Borrowers Default on Mortgages?*. *The Quarterly Journal of Economics*, page qjac040.

Gourinchas, P.-O. and Parker, J. A. (2002). Consumption Over the Life Cycle. *Econometrica*, 70(1):47–89.

Griffith, R., Leibtag, E., Leicester, A., and Nevo, A. (2009). Consumer Shopping Behavior: How Much Do Consumers Save? *Journal of Economic Perspectives*, 23(2):99–120.

Gruber, J. (1997). The Consumption Smoothing Benefits of Unemployment Insurance. *The American Economic Review*, 87(1):192–205.

Hamermesh, D. S. (1984). Consumption During Retirement: The Missing Link in the Life Cycle. *The Review of Economics and Statistics*, 66(1):1–7. Publisher: The MIT Press.

Hausman, J. and Paquette, L. (1987). Involuntary early retirement and consumption. In *Work, Health, and Income Among the Elderly*, pages 151–175.

Hurd, M. D. and Rohwedder, S. (2005). The Retirement-Consumption Puzzle: Anticipated and Actual Declines in Spending at Retirement. Publisher: RAND Corporation.

Hurd, M. D. and Rohwedder, S. (2006). Economic Well-Being at Older Ages: Income- and Consumption-Based Poverty Measures in the HRS. Technical Report w12680, National Bureau of Economic Research.

Hurst, E. (2008). The Retirement of a Consumption Puzzle. Technical Report w13789, National Bureau of Economic Research.

Iacus, S. M., King, G., and Porro, G. (2012). Causal Inference Without Balance Checking: Coarsened Exact Matching. *Political Analysis*, 20(1):1–24.

Jacobson, L. S., LaLonde, R. J., and Sullivan, D. G. (1993). Earnings Losses of Displaced Workers. *The American Economic Review*, 83(4):685–709. Publisher: American Economic Association.

Jaravel, X. (2019). The Unequal Gains from Product Innovations: Evidence from the U.S. Retail Sector. *The Quarterly Journal of Economics*, 134(2):715–783.

Kaplan, G. and Menzio, G. (2015). The Morphology of Price Dispersion. *International Economic Review*, 56(4):1165–1206. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/iere.12134.

Koustas, D. (2018). Consumption Insurance and Multiple Jobs: Evidence from Rideshare Drivers. Working Paper.

Krueger, A. B. and Mueller, A. (2010). Job search and unemployment insurance: New evidence from time use data. *Journal of Public Economics*, 94(3):298–307.

Landais, C. and Spinnewijn, J. (2021). The Value of Unemployment Insurance. *Review of Economic Studies*, forthcoming.

Nevo, A. and Wong, A. (2019). The Elasticity of Substitution Between Time and Market Goods: Evidence from the Great Recession. *International Economic Review*, 60(1):25–51. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/iere.12343.

Olafsson, A. and Pagel, M. (2018). The Retirement-Consumption Puzzle: New Evidence from Personal Finances. Technical Report w24405, National Bureau of Economic Research.

Stephens, M. and Toohey, D. (2018). Changes in Nutrient Intake at Retirement. Technical Report w24621, National Bureau of Economic Research, Cambridge, MA.

Sun, L. and Abraham, S. (2021). Estimating dynamic treatment effects in event studies with heterogeneous treatment effects. *Journal of Econometrics*, 225(2):175–199.

8 Tables

Table 1: Kilts Nielsen Consumer Panel Summary Statistics

	(1) Life-Cycle Sample			(2) Retirement Sample		
	Mean	Median	SD	Mean	Median	SD
Male Employment	0.73	1	0.45	0.46	0	0.50
Male Full-Time	0.68	1	0.47	0.37	0	0.48
Female Employment	0.60	1	0.49	0.47	0	0.50
Female Full-Time	0.48	0	0.50	0.33	0	0.47
Two Heads	0.51	1	0.50	0.72	1	0.45
No Male Head	0.28	0	0.45	0.21	0	0.41
No Female Head	0.20	0	0.40	0.074	0	0.26
Married	0.48	0	0.50	0.69	1	0.46
Household Size	2.61	2	1.46	2.03	2	0.87
Any Children Present	0.34	0	0.47	0.047	0	0.21
Male Age	49.8	50	12.7	64.6	65	6.48
Female Age	48.4	49	12.9	62.6	63	6.40
Non-Hispanic White	0.70	1	0.46	0.84	1	0.37
Black	0.12	0	0.33	0.087	0	0.28
Hispanic	0.12	0	0.33	0.037	0	0.19
Male Education	13.9	14	2.36	14.3	14	2.37
Female Education	13.9	14	2.17	14.3	14	2.14
Total Expenditure	4766.4	4106.6	3003.4	5283.4	4712.5	3035.5
Binned Household Income	73555.8	57543.8	55324.2	68912.3	57543.8	47256.8
Years in Panel	7.62	7	4.73	10.4	11	3.71
Any Retirement	0.090	0	0.29	1	1	0
Total Retirements	0.095	0	0.32	1.02	1	0.38
Any Unemployment Spell	0.097	0	0.30	0.029	0	0.17
Total Unemployment Spells	0.11	0	0.36	0.030	0	0.19
Observations	814938			107877		

Notes: This table presents summary statistics for the three main analysis samples. Male always refers to the male household head, female refers to the female household head. Education is expressed in years. Household income is deflated by the PCE, total expenditures are deflated by the CPI Food at Home for Urban Consumer; both are expressed in 2012 dollars. Binned household income in the Nielsen Consumer Panel is translated into dollars based on the mean income in the same income bin the IRS Statistics of Income database. For the life-cycle sample, all values are weighted using projection factors to make sample representative of US population. For the retirement sample, all values are weighted by the inverse of years in the sample, so that each household has the same weight. Total observations correspond to household years.

Table 2: Summary Statistics Health and Retirement Survey Data

	Mean	Median	SD
Panel A: Not Retired			
Household Size	2.5668	2	1.3633
Married	0.6293	1	0.4830
Age	58.201	57	8.1996
White	0.7047	1	0.4562
Black	0.1759	0	0.3808
Hispanic	0.1572	0	0.3640
Years of Education	13.092	13	3.2349
Log Household Income (2012 USD)	10.815	11.098	1.8031
Log Total Expenditure (2012 USD)	10.597	10.621	0.7381
Log Nondurable Expenditure (2012 USD)	9.9373	9.9635	0.7508
Weekly Hours Work for Pay	30.687	40	20.492
Weekly Hours Homeproduction	25.485	20	23.740
N	19067		
Panel B: Retired			
Household Size	2.0645	2	1.0727
Married	0.5460	1	0.4979
Age	71.207	71	9.7238
White	0.7873	1	0.4092
Black	0.1562	0	0.3631
Hispanic	0.08608	0	0.2805
Years of Education	12.564	12	3.0480
Log Household Income (2012 USD)	10.367	10.443	1.3565
Log Total Expenditure (2012 USD)	10.337	10.354	0.7762
Log Nondurable Expenditure (2012 USD)	9.7941	9.8257	0.8073
Weekly Hours Work for Pay	3.5184	0	10.181
Weekly Hours Home Production	26.457	21	24.009
N	42892		
Panel C: Full Sample			
Household Size	2.2191	2	1.1926
Married	0.5717	1	0.4948
Age	67.204	66	11.053
White	0.7618	1	0.4260
Black	0.1623	0	0.3688
Hispanic	0.1079	0	0.3103
Years of Education	12.726	12	3.1161
Log Household Income (2012 USD)	10.505	10.626	1.5222
Log Total Expenditure (2012 USD)	10.411	10.430	0.7744
Log Nondurable Expenditure (2012 USD)	9.8350	9.8646	0.7942
Weekly Hours Work for Pay	11.655	0	18.794
Weekly Hours Home Production	26.166	20	23.933
N	62664		

Notes: This table presents summary statistics for the HRS CAMS sample, broken out by retired and not-yet-retired households.

Table 3: Distribution of the Effort and Quality Distributions

Mean	Price Effects	Bulk Effects	Price + Bulk Effects	Quality Adjustments
SD	7.97	5.76	9.93	11.93
p5	-13.14	-9.51	-15.17	-19.96
p10	-8.66	-7.09	-11.56	-15.01
p25	-3.45	-3.44	-6.03	-7.36
p50	0.38	0.1	-0.28	0.35
p75	4.15	3.47	5.79	7.62
p90	8.49	6.73	11.95	14.27
p95	11.48	8.97	16.12	18.63

Notes: This table presents moments for my measure of price effects, bulk effects, combined price and bulk effects, as well as the quality adjustment metric. These statistics correspond to the baseline sample of the data presented in Table 1.

Table 4: Validating the Quality Measures

	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: National Quality						
Log Household Income	0.0643	0.0231	0.0523	0.0609	0.0564	0.0192
	(0.0005)	(0.0007)	(0.0006)	(0.0006)	(0.0006)	(0.0009)
Lifetime Income Rank		0.1375				0.1010
		(0.0021)				(0.0026)
Log Home Price			0.0463			0.0404
			(0.0007)			(0.0008)
Imputed Wealth Ventile				0.1657		0.0245
				(0.0060)		(0.0061)
Highschool					1.3839	1.9243
					(0.3287)	(0.4070)
Some College					3.0698	3.0269
					(0.3272)	(0.4060)
College Grad					4.6408	3.8905
					(0.3292)	(0.4087)
Postgrad					6.1632	4.7137
					(0.3394)	(0.4205)
Panel B: Local Quality						
Log Household Income	0.0572	0.0220	0.0537	0.0544	0.0502	0.0188
	(0.0006)	(0.0007)	(0.0007)	(0.0006)	(0.0006)	(0.0009)
Lifetime Income Rank		0.1177				0.1043
		(0.0022)				(0.0028)
Log Home Price			0.0133			0.0069
			(0.0008)			(0.0009)
Imputed Wealth Ventile				0.1507		0.0402
				(0.0062)		(0.0067)
Highschool					1.1226	1.7712
					(0.3293)	(0.4123)
Some College					2.5226	2.8361
					(0.3274)	(0.4109)
College Grad					3.8805	3.7851
					(0.3299)	(0.4147)
Postgrad					5.3974	4.8033
					(0.3414)	(0.4291)
Observations	790983	770597	615243	598557	790983	453658

Notes: This table reports regressions of the national and local quality measure on six separate measures of socioeconomic status: Current income (in 2012 USD), lifetime income percentile, log real house prices from Zillow (for households living in single-family homes or condos/coops), imputed household wealth from the Survey of Consumer Finances, and indicators for educational attainment with no high school degree as the reference category. All regressions control for household size, household composition (two household heads, married, otherwise related household heads, a single male or female head), and the age and presence of children. The estimation sample for each regression is the baseline sample from Table 1. Clustered standard errors in parentheses.

Table 5: Reported MPCs at Retirement

	20% Income Increase			20% Income Decline		
	MPC (pp.)	ln MPC	MPC $\neq 0$	MPC (pp.)	ln MPC	MPC $\neq 0$
DID Estimate	7.512	0.258	0.085	-1.440	-0.096	0.012
	(1.882)	(0.090)	(0.026)	(2.140)	(0.045)	(0.015)
N	2850	1234	2850	3113	2646	3113

Notes: This table reports difference-in-difference coefficients from a regression of reported MPCs out of a hypothetical income increase (decline) just before and just after retirement. Estimates are estimated following de Chaisemartin and D'Haultfœuille (2020) and are robust to treatment effect heterogeneity. Bootstrap standard errors clustered at the household level in parentheses.

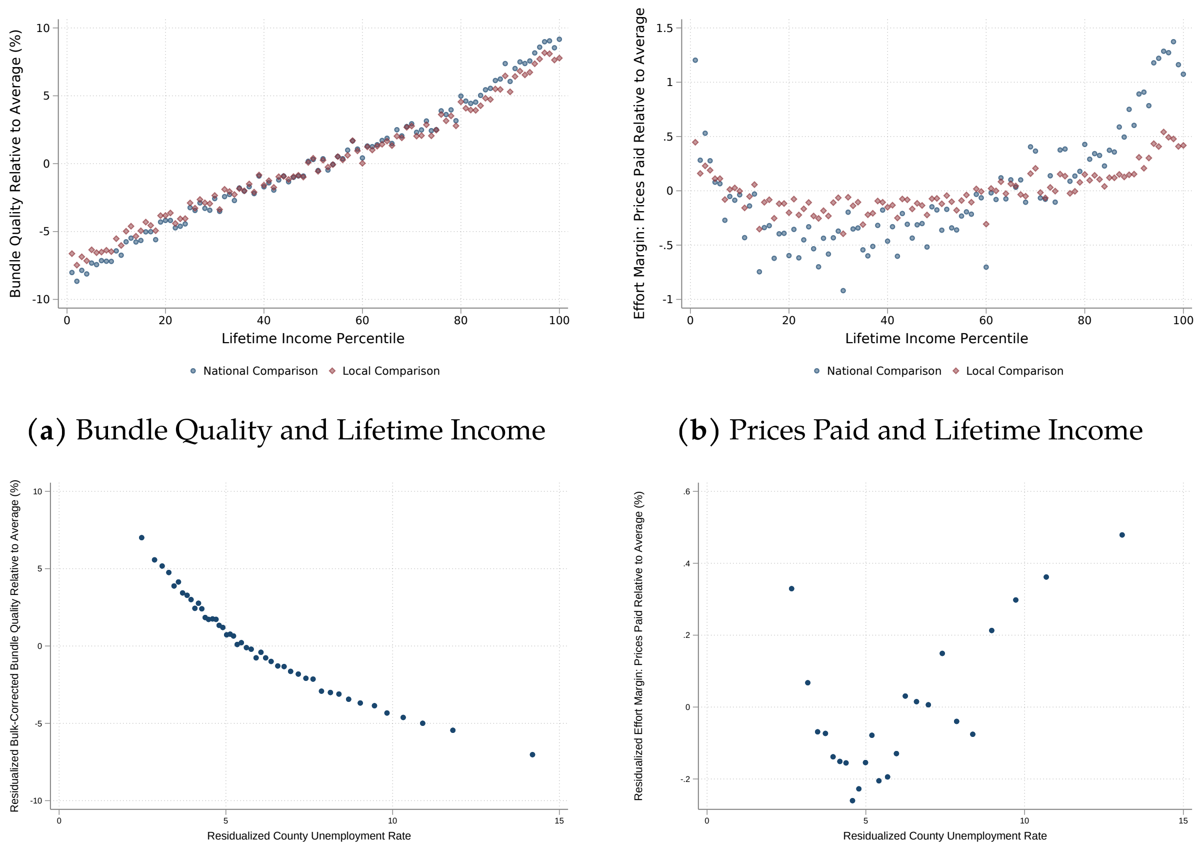
Table 6: Shopping Behavior and Employment

	(1) Prices	(2) Prices	(3) Prices	(4) Quality	(5) Quality	(6) Quality
			Panel A: Households Ages 25-54			
Male Part Time	0.190	0.234	-0.0482	-2.069***	1.830***	-0.315
	(0.158)	(0.158)	(0.131)	(0.335)	(0.315)	(0.229)
Male Not Employed	0.0533	0.152	0.0354	-5.225***	-0.211	-0.500**
	(0.109)	(0.109)	(0.117)	(0.198)	(0.194)	(0.176)
Female Part Time	-0.913***	-0.775***	-0.193*	-1.751***	0.472**	-0.346*
	(0.0900)	(0.0899)	(0.0788)	(0.162)	(0.151)	(0.137)
Female Not Employed	-0.922***	-0.762***	-0.478***	-3.212***	-0.0605	-0.630***
	(0.0733)	(0.0759)	(0.0823)	(0.137)	(0.129)	(0.141)
Income Controls		X			X	
Household FE			X			X
N	286923	286923	286923	286919	286919	286919
			Panel B: Households Ages 55-74			
Male Part Time	-0.390**	-0.216	-0.0505	-3.056***	-0.421	-0.224
	(0.142)	(0.142)	(0.0818)	(0.351)	(0.294)	(0.176)
Male Not Employed	-0.520***	-0.309**	-0.0769	-3.887***	-0.536**	-0.645***
	(0.106)	(0.105)	(0.0698)	(0.197)	(0.186)	(0.121)
Female Part Time	-1.019***	-0.904***	-0.207**	-2.125***	-0.297	-0.607***
	(0.139)	(0.137)	(0.0786)	(0.263)	(0.246)	(0.147)
Female Not Employed	-0.698***	-0.571***	-0.191*	-2.790***	-0.423*	-0.751***
	(0.0970)	(0.0967)	(0.0795)	(0.198)	(0.185)	(0.144)
Income Controls		X			X	
Household FE			X			X
N	247646	247646	247646	247646	247646	247646

Notes: This paper reports the effect of work status of male and female household heads on household level prices and household bundle quality. The underlying sample is the baseline sample in Table 1, restricted to households with two household heads. Household age is defined as the average age of the two household heads, rounded to the nearest integer. The omitted category in both cases is a household head working full time (at least 30 hours a week). All regressions control for household size, the relationship between the household heads, indicators for the age and presence of children and year fixed effects. Regressions are weighted to be representative of the US population. Clustered standard errors in parentheses.

9 Figures

Figure 1: Validating the Quality Measure



(**a**) Bundle Quality and Lifetime Income

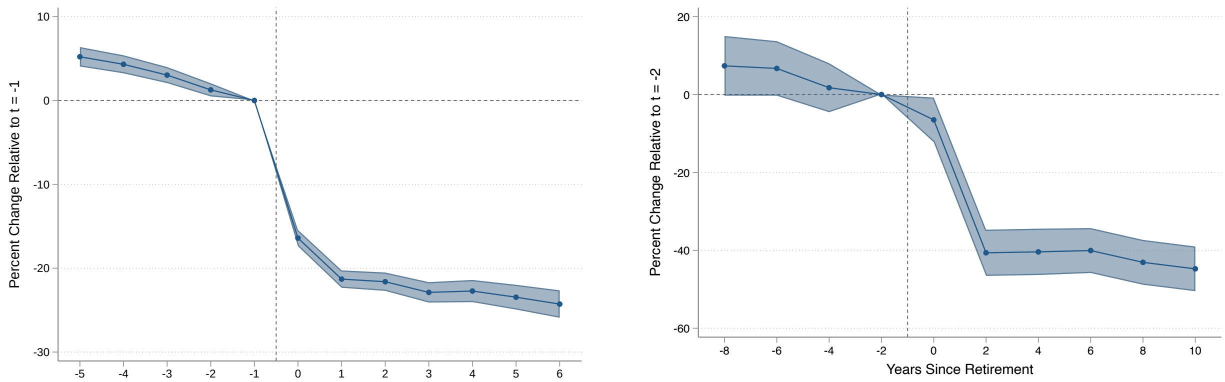
(**b**) Prices Paid and Lifetime Income

(**c**) Bundle Quality and Local Unemployment

(**d**) Prices Paid and Local Unemployment

Notes: This figure presents two validation exercises for disentangling price, bulk, and quality effects. The effort and quality margins are as defined in Equations 7 and 10. The underlying data for all figures is the baseline sample of Table 1.

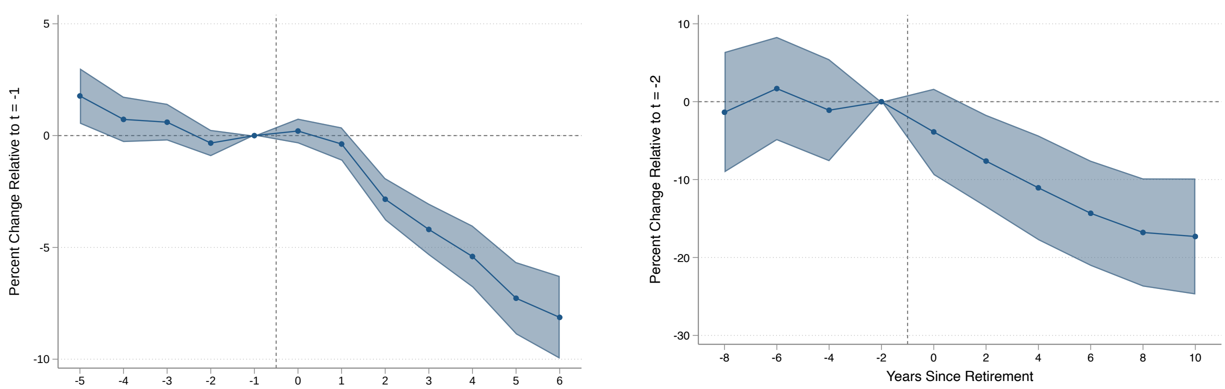
Figure 2: The Path of Income Around Retirement



(**a**) Scanner Data (**b**) Consumption Survey

Notes: This figure presents the path of income around retirement in the Kilts Nielsen Consumer Panel and the Health and Retirement Survey estimated according to Equation 11. The underlying data for Panel (a) is the retirement sample of Table 1, for Panel (b) is the retirement sample of Table 2. Standard errors are clustered at the household level.

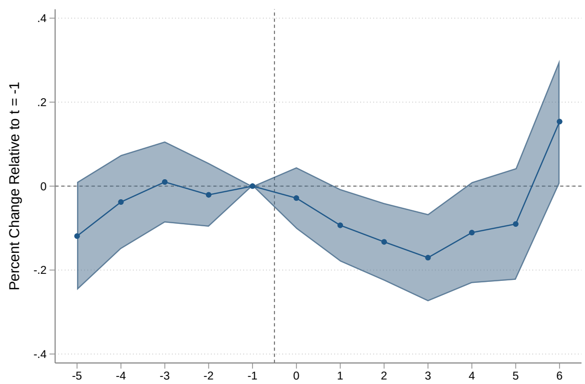
Figure 3: The Path of Expenditure Around Retirement



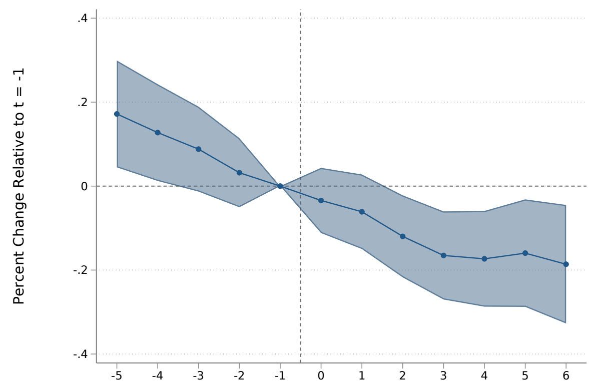
(**a**) Scanner Data (**b**) Consumption Survey

Notes: This figure presents the path of expenditure around retirement in the Kilts Nielsen Consumer Panel and in the Health and Retirement Survey estimated according to Equation 11. Expenditure in the KNCP is defined as spending on non-magnet data product modules covered in all panel years. Expenditure in the HRS is total reported expenditure on non-durables (excluding housing). The underlying data for Panel (a) is the retirement sample of Table 1, for Panel (b) is the retirement sample of Table 2. Standard errors are clustered at the household level.

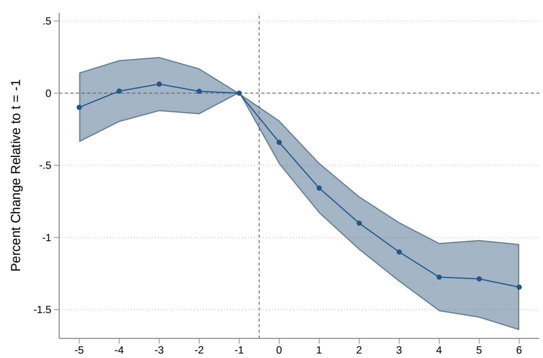
Figure 4: Shopping Adjustments Around Retirement



(**a**) Prices Paid



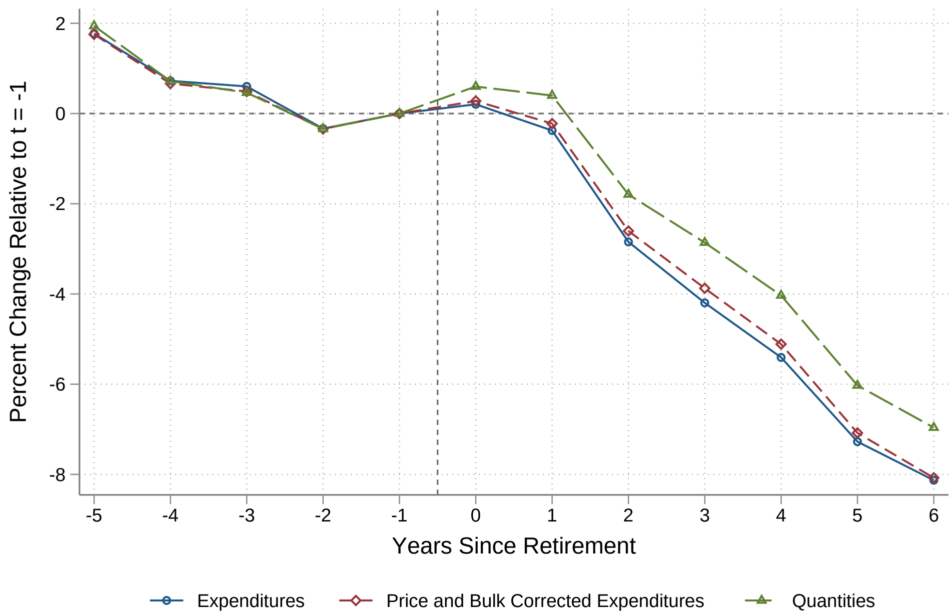
(**b**) Cost Relative to Average Bulk Allocation



(**c**) Bundle Quality

Notes: This figure presents the evolution of price effects, bulk effects, and quality adjustments estimated according to Equation 11. The underlying data for each panel is the retirement sample of Table 1. Standard errors are clustered at the household level.

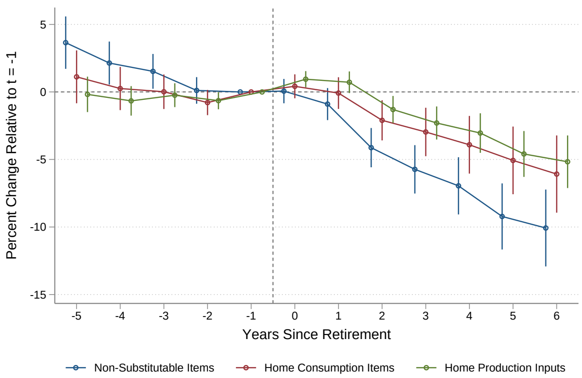
Figure 5: Decomposing the Post-Retirement Expenditure Drop



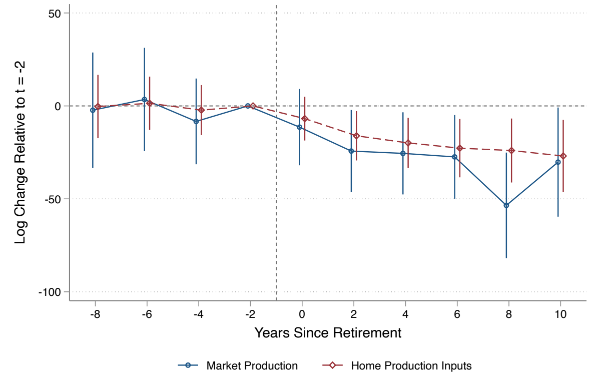
(**a**) Expenditures and Corrected Expenditures

Notes: This figure presents the evolution of price effects, bulk effects, and quality adjustments estimated according to Equation 11 as well as uncorrected expenditure, expenditure at average prices and the average bulk allocation, and expenditure at average prices, the average bulk allocation, and average within-product module quality. The underlying data for each panel is the retirement sample of Table 1.

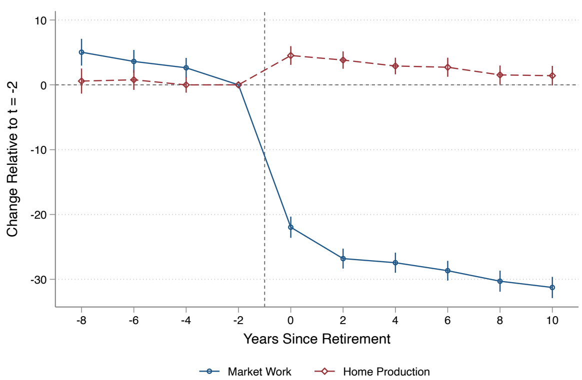
Figure 6: Home Production, Home Consumption and Expenditure Timing in Retirement



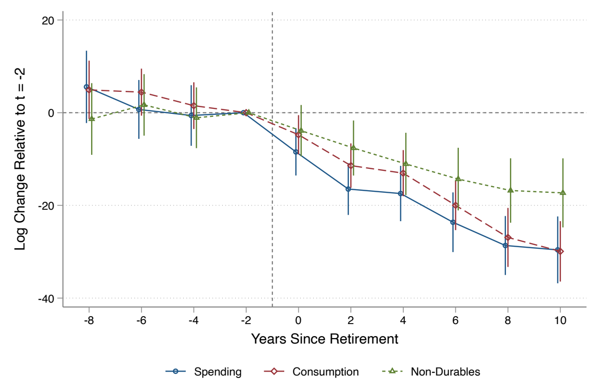
(**a**) Subsets of Spending by Substitutability



(**b**) Market and Home Production Expenditure



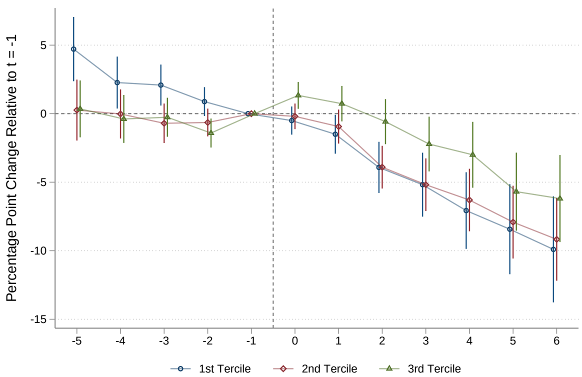
(**c**) Time Working vs. in Home Production



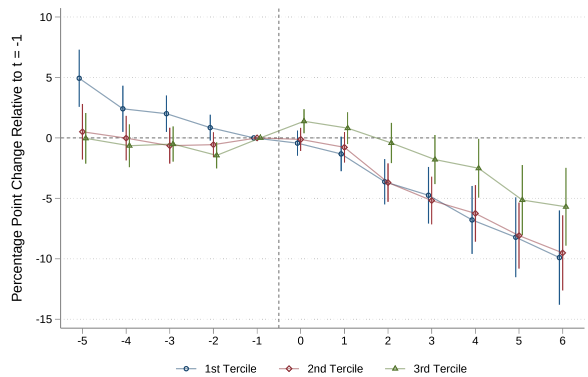
(**d**) Total vs. Consumption Expenditure

Notes: This figure presents the path of expenditure categories around retirement in the Kilts Nielsen Consumer Panel as well as expenditure categories and time use in the Health and Retirement Survey estimated according to Equation 11. Non-substitutable items are non-food grocery items, health and beauty items, and general merchandise. Home consumption items are prepared foods and deli items, food items that do not require preparation, and alcohol. Home production inputs are dry grocery items, packaged meat, fresh produce, frozen meat and vegetables, and dairy products. Home production items in the HRS are food at home, cleaning products, gardening products expenditures. Market production items are their immediate market good substitutes (food away from home, cleaning services, and gardening services). Home production time is time spent cleaning, doing gardening work, preparing meals and cleaning up and time spent shopping or running errands. The underlying data for Panel (a) is the retirement sample of Table 1; the data for Panels (b)-(d) is the retirement sample of Table 2. Standard errors are clustered at the household level.

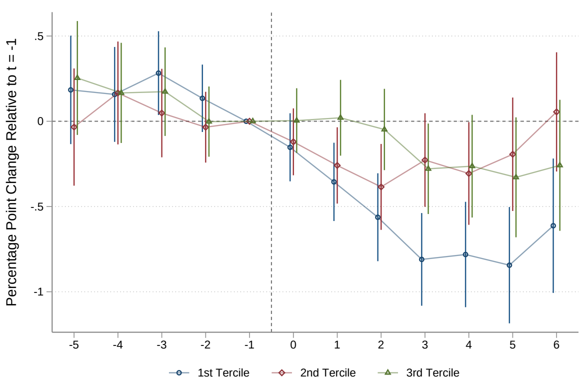
Figure 7: Heterogeneity by Size of Income Drop



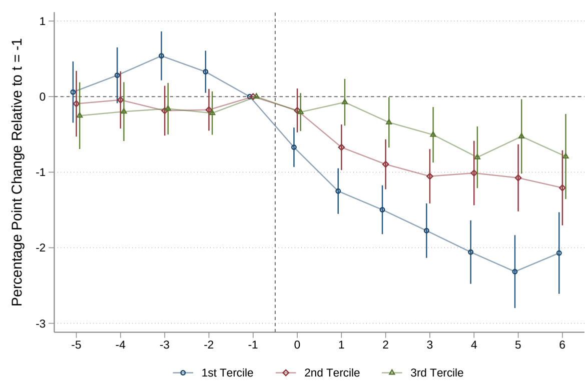
(**a**) Expenditures



(**b**) Price & Bulk Corrected Expenditures



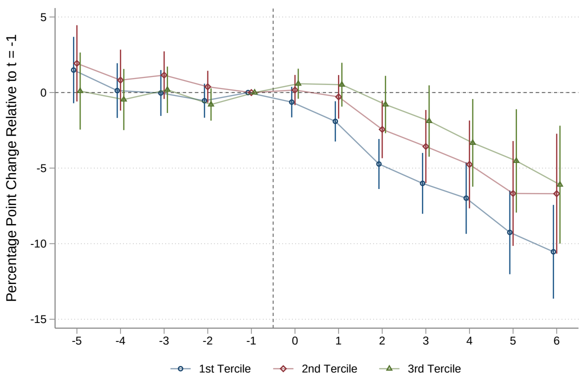
(**c**) Price and Bulk Effects



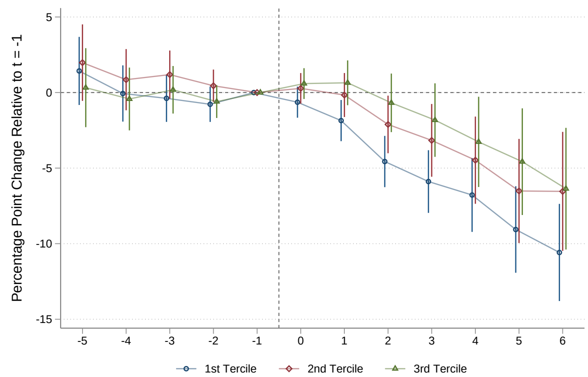
(**d**) Quality Adjustments

Notes: This figure presents the evolution of expenditures, price and bulk corrected expenditures, price and bulk effects, and quality adjustments for each tercile of the post-retirement income drop. The income drop is defined as the log difference in the average income over the three years preceding and immediately succeeding retirement. The underlying data for each panel is the retirement sample of Table 1. Standard errors are clustered at the household level.

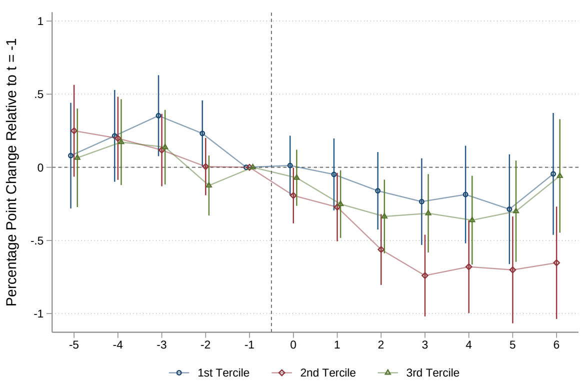
Figure 8: Heterogeneity by Pre-Retirement Wealth



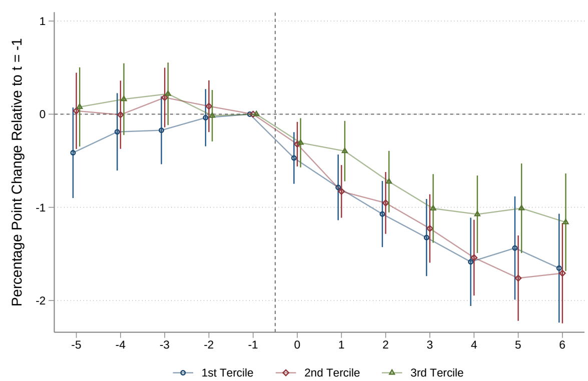
(**a**) Expenditures



(**b**) Price & Bulk Corrected Expenditures



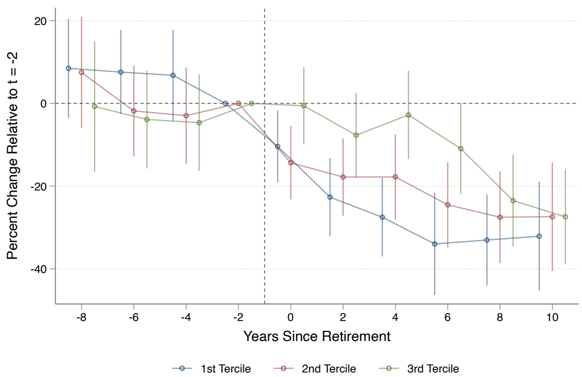
(**c**) Price and Bulk Effects



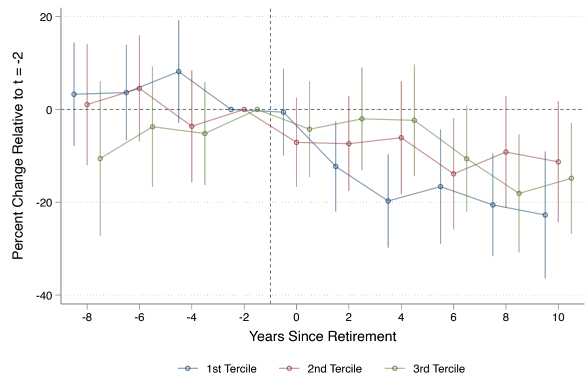
(**d**) Quality Adjustments

Notes: This figure presents the evolution of expenditures, price and bulk corrected expenditures, price and bulk effects, and quality adjustments for each tercile of pre-retirement household wealth, where household wealth is a cell based match to the Survey of Consumer Finances. The underlying data for each panel is the retirement sample of Table 1. Standard errors are clustered at the household level.

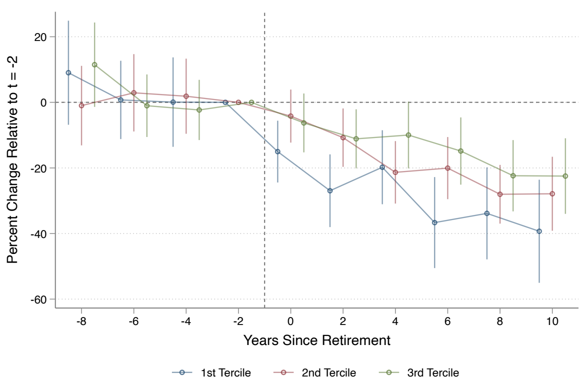
Figure 9: Heterogeneity of Broader Expenditures by Income and Wealth



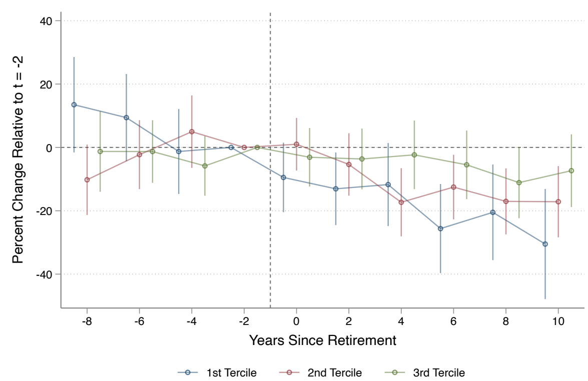
(**a**) By Income Drop: Expenditure



(**b**) By Income Drop: Non-Durables



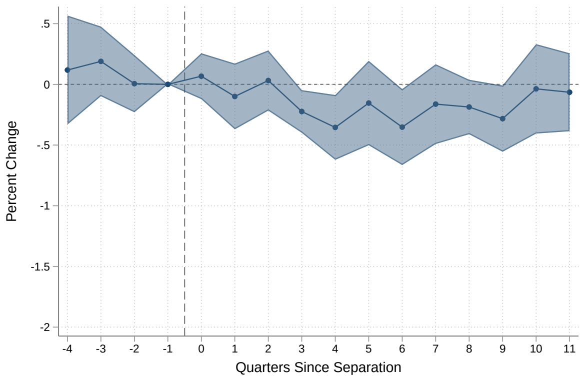
(**c**) By Wealth-to-Income: Expenditure



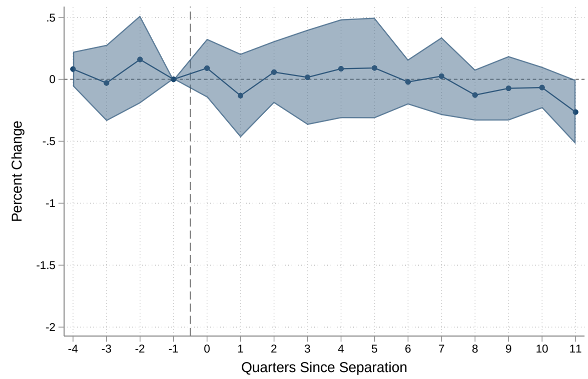
(**d**) By Wealth-to-Income: Non-Durables

Notes: This figure presents the evolution of total expenditure and non-durable expenditure by the size of the post-retirement income drop and the pre-retirement wealth-to-income ratios. The underlying data for each panel is the retirement sample of Table 2.

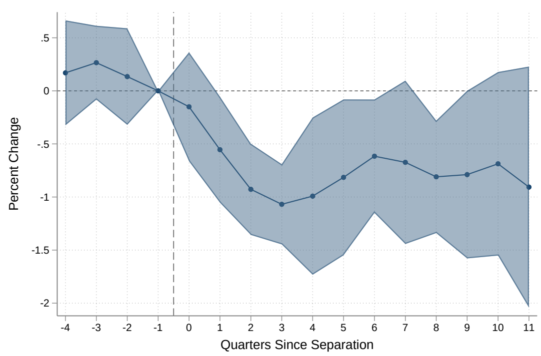
Figure 10: Shopping Adjustments Around Nonemployment



(**a**) Prices Paid



(**b**) Cost Relative to Average Bulk Allocation



(**c**) Bundle Quality

Notes: This figure presents the evolution of price effects, bulk effects, and quality adjustments estimated according to Equation 11 for households undergoing a non-employment spell lasting at most two years. Households heads are reporting to be employed at $t = -1$, report non-employment at $t = 3$, and are reporting being employed again in either $t = 7$ or $t = 11$. Standard errors are clustered at the household level.

A Decomposing Expenditure

Since we are interested in decomposing quantities, prices, bulk effects, and the quality of the consumption bundle, we can proceed as follows. Fixing a "bundle" at the product-module level, let p denote prices at the item level, b denote bulk allocations, and μ denote quality, we have

$$e_i = \text{Cost of Bundle at } p_i,\, b_i,\, \mu_i$$

which we can express as

$$= \text{Cost of of Bundle at } \bar{p}_i,\, \bar{b}_i,\, \bar{\mu}_i \cdot \frac{\text{Cost of Bundle at } p_i,\, b_i,\, \mu_i}{\text{Cost of of Bundle at } \bar{p}_i,\, \bar{b}_i,\, \bar{\mu}_i}$$

This fixes the bundle at the product-module level. A household with a given bundle can pay more or less than the average by purchasing varieties of higher (or lower) quality, by paying more (or less) for a given variety k or by buying more (or less) in bulk. Taking logs, we have:

$$\begin{aligned}
\ln e_i &= \ln\left(\text{Cost of of Bundle at } \bar{p}_i,\, \bar{b}_i,\, \bar{\mu}_i\right) \\
&\quad + \ln\left(\text{Cost of Bundle at } p_i,\, b_i,\, \mu_i\right) - \ln\left(\text{Cost of of Bundle at } \bar{p}_i,\, \bar{b}_i,\, \bar{\mu}_i\right)
\end{aligned}$$

Below, I show that the deviation of the cost of the bundle deviation from the cost of the bundle at average barcode-level prices, the average bulk allocation, and average quality is additively separable in each of the sub-components. Therefore, we have:

$$\begin{aligned}
\ln e_i &= \ln\left(\text{Cost of of Bundle at } \bar{p}_i,\, \bar{b}_i,\, \bar{\mu}_i\right) \\
&\quad + \ln\left(\text{Cost of Bundle at } p_i\right) - \ln\left(\text{Cost of Bundle at } \bar{p}_i\right) \\
&\quad + \ln\left(\text{Cost of Bundle at } b_i\right) - \ln\left(\text{Cost of Bundle at } \bar{b}_i\right) \\
&\quad + \ln\left(\text{Cost of Bundle at } \mu_i\right) - \ln\left(\text{Cost of Bundle at } \bar{\mu}_i\right) \\
&= \text{Quantity Effects}_i + \text{Price Effects}_i + \text{Bulk Effects}_i + \text{Quality Effects}_i
\end{aligned}$$

We can decompose expenditure into quantity, quality, price, and bulk effects as follows. Starting with expenditure, we have:

$$\text{Expenditure}_{i,t} = \sum_k P_{i,t,k} Q_{i,t,k}$$

Denoting the within-m average price by $\bar{P}_{m,t}$, we can decompose this into the actual bundle at product-module average prices and the ratio of the actual cost of the bundle and the cost of the bundle at module-average prices:

$$= \left(\sum_{m} \sum_{k \in m} \bar{P}_{t,m} Q_{i,t,k} \right) \cdot \frac{\sum_k P_{i,t,k} Q_{i,t,k}}{\sum_m \sum_{k \in m} \bar{P}_{t,m} Q_{i,t,k}}$$

Taking logs, we have

$$\ln \text{Expenditure}_{i,t} = \ln \left(\sum_{m} \sum_{k \in m} \bar{P}_{i,t,m} Q_{i,t,k} \right) + \ln \left(\frac{\sum_k P_{i,t,k} Q_{i,t,k}}{\sum_m \sum_{k \in m} \bar{P}_{i,t,m} Q_{i,t,k}} \right)$$

Importantly, the first term captures changes in quantities only (either increasing $Q_{i,t,k}$ for fixed k or switching the bundle between k and l for $k \neq l$). We can think of the last term as an aggregate shopping margin that is captures the log difference between the actual cost of the bundle household i decided to buy and cost of the same product-module-level quantities at product-module-level mean prices. We can decompose this adjustment margin as follows:

$$\textit{Aggregate Margin}_{i,t} = \ln \left(\sum_{k} P_{i,t,k} \cdot Q_{i,t,k} \right) - \ln \left(\sum_{m} \left(\frac{\sum_{k \in m} \sum_j P_{j,t,k} Q_{j,t,k}}{\sum_{k \in m} \sum_j Q_{j,t,k}} \cdot \sum_{k \in m} Q_{i,t,k} \right) \right)$$

Summing over the products of actual quantities and average prices within a product module first and then aggregating over product modules will yield the same total as directly summing over the products of actual quantities and average prices over all UPCs. Therefore, we can add and substract the actual bundle at average prices and get

$$\begin{aligned} &= \ln \left(\sum_{k} P_{i,t,k} \cdot Q_{i,t,k} \right) - \ln \left(\sum_{m} \left(\frac{\sum_{k \in m} \sum_j P_{j,t,k} Q_{j,t,k}}{\sum_{k \in m} \sum_j Q_{j,t,k}} \cdot \sum_{k \in m} Q_{i,t,k} \right) \right) \\ &+ \ln \left(\sum_{m} \sum_{k \in m} \frac{\sum_j P_{j,t,k} Q_{j,t,k}}{\sum_j Q_{j,t,k}} \cdot Q_{i,t,k} \right) - \ln \left(\sum_{k} \frac{\sum_j P_{j,t,k} Q_{j,t,k}}{\sum_j Q_{j,t,k}} \cdot Q_{i,t,k} \right) \end{aligned}$$

We can re-arrange this to yield

$$\begin{aligned} &= \ln \left(\sum_{k} P_{i,t,k} \cdot Q_{i,t,k} \right) - \ln \left(\sum_{k} \frac{\sum_j P_{j,t,k} Q_{j,t,k}}{\sum_j Q_{j,t,k}} \cdot Q_{i,t,k} \right) \\ &+ \ln \left(\sum_{m} \sum_{k \in m} \frac{\sum_j P_{j,t,k} Q_{j,t,k}}{\sum_j Q_{j,t,k}} \cdot Q_{i,t,k} \right) \end{aligned}$$

$$
\begin{aligned}
&- \ln\left(\sum_m \left(\frac{\sum_{k\in m}\sum_j P_{j,t,k}Q_{j,t,k}}{\sum_{k\in m}\sum_j Q_{j,t,k}} \cdot \sum_{k\in m} Q_{i,t,k}\right)\right) \\
&= \textit{Effort Margin}_{i,t} + \textit{Uncorrected Quality Margin}_{i,t}
\end{aligned}
$$

Turning to the uncorrected quality margin, we can follow Griffith et al. (2009) and we break out each product module into five quintiles of the size distribution. Then, we can define UPCs in the top two quintiles (or the Top 40% of the within-product-module size distribution) as bulk items.[18] We can now add and subtract the cost of the quantities the household actually bought for a given product-module-by-bulk level but at the product-module-by-bulk mean price:

$$
\begin{aligned}
&= \ln\left(\sum_m\sum_k \frac{\sum_j P_{j,t,k}Q_{j,t,k}}{\sum_j Q_{j,t,k}} \cdot Q_{i,t,k}\right) - \ln\left(\sum_m \left(\frac{\sum_k\sum_j P_{j,t,k}Q_{j,t,k}}{\sum_k\sum_j Q_{j,t,k}} \cdot \sum_k Q_{i,t,k}\right)\right) \\
&+ \ln\left(\sum_{m\times b} \left(\frac{\sum_{k\in m\times b}\sum_j P_{j,t,k}Q_{j,t,k}}{\sum_{k\in m\times b}\sum_j Q_{j,t,k}} \cdot \sum_{k\in m\times b} Q_{i,t,k}\right)\right) \\
&- \ln\left(\sum_{m\times b} \left(\frac{\sum_{k\in m\times b}\sum_j P_{j,t,k}Q_{j,t,k}}{\sum_{k\in m\times b}\sum_j Q_{j,t,k}} \cdot \sum_{k\in m\times b} Q_{i,t,k}\right)\right)
\end{aligned}
$$

As before, we can rearrange this to get:

$$
\begin{aligned}
&= \ln\left(\sum_m\sum_k \frac{\sum_j P_{j,t,k}Q_{j,t,k}}{\sum_j Q_{j,t,k}} \cdot Q_{i,t,k}\right) - \ln\left(\sum_{m\times b} \left(\frac{\sum_{k\in m\times b}\sum_j P_{j,t,k}Q_{j,t,k}}{\sum_{k\in m\times b}\sum_j Q_{j,t,k}} \cdot \sum_{k\in m\times b} Q_{i,t,k}\right)\right) \\
&+ \ln\left(\sum_{m\times b} \left(\frac{\sum_{k\in m\times b}\sum_j P_{j,t,k}Q_{j,t,k}}{\sum_{k\in m\times b}\sum_j Q_{j,t,k}} \cdot \sum_{k\in m\times b} Q_{i,t,k}\right)\right) \\
&- \ln\left(\sum_m \left(\frac{\sum_k\sum_j P_{j,t,k}Q_{j,t,k}}{\sum_k\sum_j Q_{j,t,k}} \cdot \sum_k Q_{i,t,k}\right)\right) \\
&= \textit{Quality Margin}_{i,t} + \textit{Bulk Margin}_{i,t}
\end{aligned}
$$

Finally, we can put all of these pieces together and get:

$$
\begin{aligned}
\ln \text{Expenditure}_{i,t} &= \ln\left(\sum_m\sum_{k\in m} \bar{P}_{i,t,m}Q_{i,t,k}\right) - \textit{Quality Margin}_{i,t} - \textit{Effort Margin}_{i,t} - \textit{Bulk Margin}_{i,t} \\
&= \textit{Quantity Effects}_{i,t} - \textit{Quality Margin}_{i,t} - \textit{Effort Margin}_{i,t} - \textit{Bulk Margin}_{i,t}
\end{aligned}
$$

[18] Note that I use *total quantities* to define item size. For example, a twelve-pack of cans of Coca-Cola will be defined as equaling 144 fluid ounces (12 times 12 fluid ounces), roughly similar to two-liter bottles of Coca-Cola (135.2 fluid ounces).

Intuitively, households can bring down the price of their product-module level bundle by paying less for the same UPC, by buying more in bulk, or by substituting towards lower quality items within a product module. In practice, not all UPCs in the KNCP data have interpretable quality information. For example, some households collect "magnet data" which refers to items that are sold by weight and do not have an associated UPC (e.g., fresh meat at the counter or by-weight produce). Households record these purchases in counts (instead of weight or volume), which makes interpreting the quantity for these items impossible. Therefore, I drop magnet data from my analysis. Another issue are durables and semi-durables recorded in "General Merchandise". While most goods in this category have their own UPC (allowing me to compare prices paid for any specific item), it is unclear to what extent price differences within a product module will reflect quality differences or differences in some other attributes. For example, one such product module is calendars. While I can observe exactly how much any one household paid for a given calendar (at the UPC level), it is unclear to what extent all price differences truly reflect quality differences in the same narrow sense as this is true in the other levels of consumption I observe. A calendar might be bigger or smaller (e.g., have an individual page for each day or a page for every workweek) so that this isn't the same narrow comparison I make elsewhere. Therefore, I keep these purchases when estimating my main shopping effort measure but drop them for my quality comparisons. For robustness, I also compute the shopping effort measure after excluding these purchases.

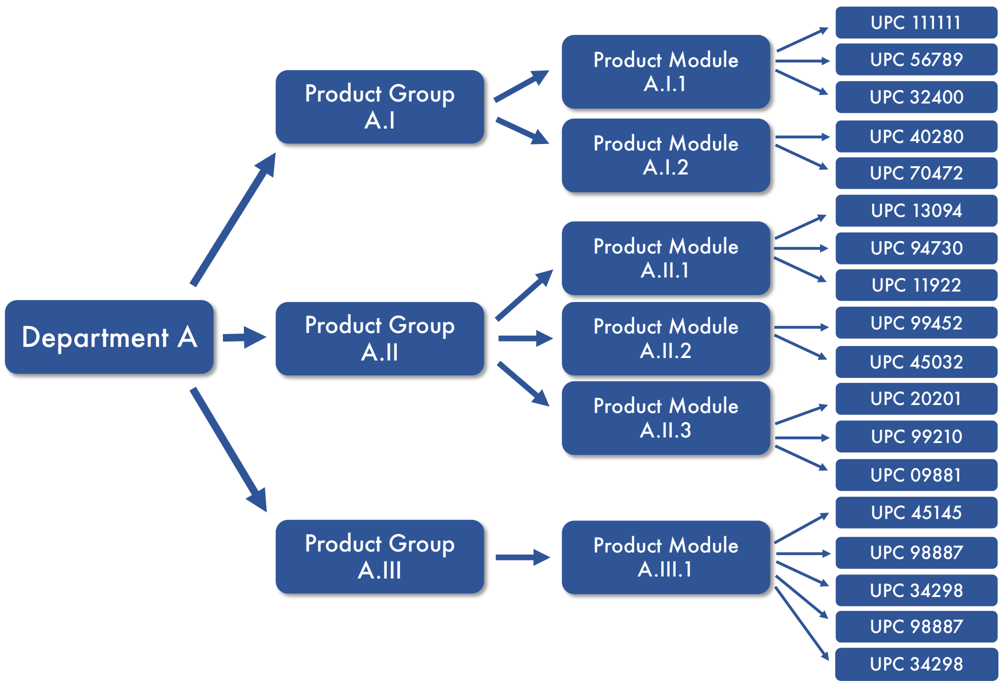
B Data Definitions and Details on Sample Construction

B.1 Matching the Nielsen Panel to the SCF

When matching information on asset holdings from the Survey of Consumer Finances to the Nielsen Panel, I follow the following procedure: For each iteration of the SCF from 2004 to 2019, I restrict the Nielsen Panel to the years surrounding the year the SCF was conducted (so I use KNCP information elicited in 2003, 2004, and 2005 to match to the 2004 SCF). I then match exactly on household structure (a cohabitating couple, a single female head, a single male head), household race (using the household head's race in the SCF and the reported "household race" from the Nielsen Panel), and on the employment status of the household head (employed or not employed). I the employ the coarsened exact matching algorithm by Iacus et al. (2012) to match on educational attainment of the household head (less than high school, high school, some college, a college degree, a post-graduate degree), income quintile (using the distribution of income in the Nielsen Panel), the number of children (no children, one child, two or more children), and the age of the household head.

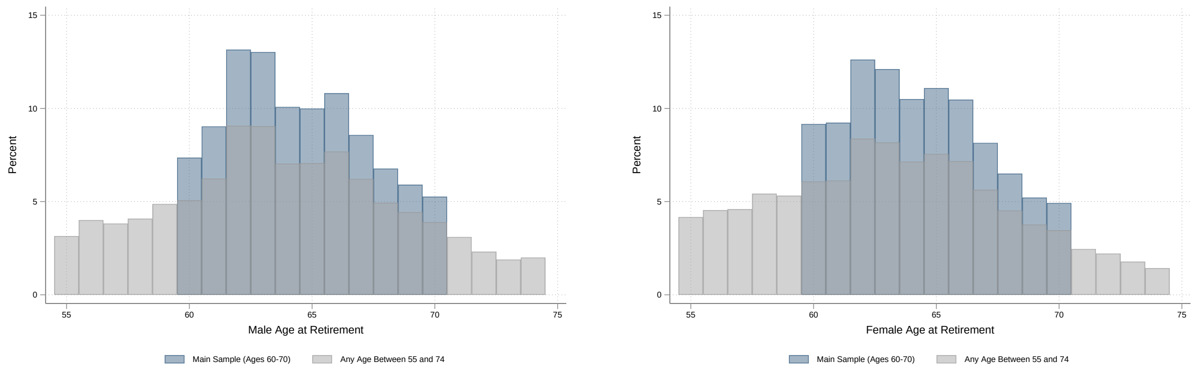
C Appendix Figures

Figure A.1: Product Hierarchy in Nielsen



Notes: An illustration of the product hierarchy in Nielsen. Each of nine departments is organized into various product groups (a total of 110), which in turn are disaggregated into product modules (a total of 1,298). Each product group may nest one or more product modules and each product module will correspond to a large number of different UPCs.

Figure A.2: Distribution of Age at Retirement in Nielsen



(**a**) Men (**b**) Women

Figure A.3: An Illustration of UPCs capturing different items

Notes: An illustration of three different sets of UPCs that amount to an exactly identical quantity of an identical good.

Figure A.4: Illustration of the Effort Margin

(**a**) Store-Brand Eggs: $2.77

(**b**) Store-Brand Eggs: $3.80

(**c**) Name-Brand Pasture-Raised Eggs: $5.99

(**d**) Name-Brand Pasture-Raised Eggs: $7.49

Notes: An illustration of the effort margin. This figure shows two different products with different prices at different points in time or at different retailers. Shopping effort refers to buying a *given* UPC at a lower price than the average for that UPC. Prices are actual prices at specific stores in Alameda County and *not* taken from the Kilts Nielsen Nielsen Panel.

Figure A.5: Illustration of the Bulk Margin

(**a**) 6-Count Eggs: $0.332 each

(**b**) 18-Count Eggs: $0.229 each

(**c**) 36-Count Eggs: $0.226 each

(**d**) 60-Count Eggs: $0.202 each

Notes: An illustration of the bulk margin. This figure shows the prices of four distinct UPCs within the product module "Eggs" with different degrees of "bulkiness". Bulk savings means purchasing eggs in larger quantities in order to reduce per-unit prices (abstracting away from realized prices). The presented prices are actual prices at specific stores in Alameda County and *not* taken from the Kilts Nielsen Nielsen Panel.

Figure A.6: Illustration of the Quality Margin

(**a**) Store-Brand Eggs: $2.77

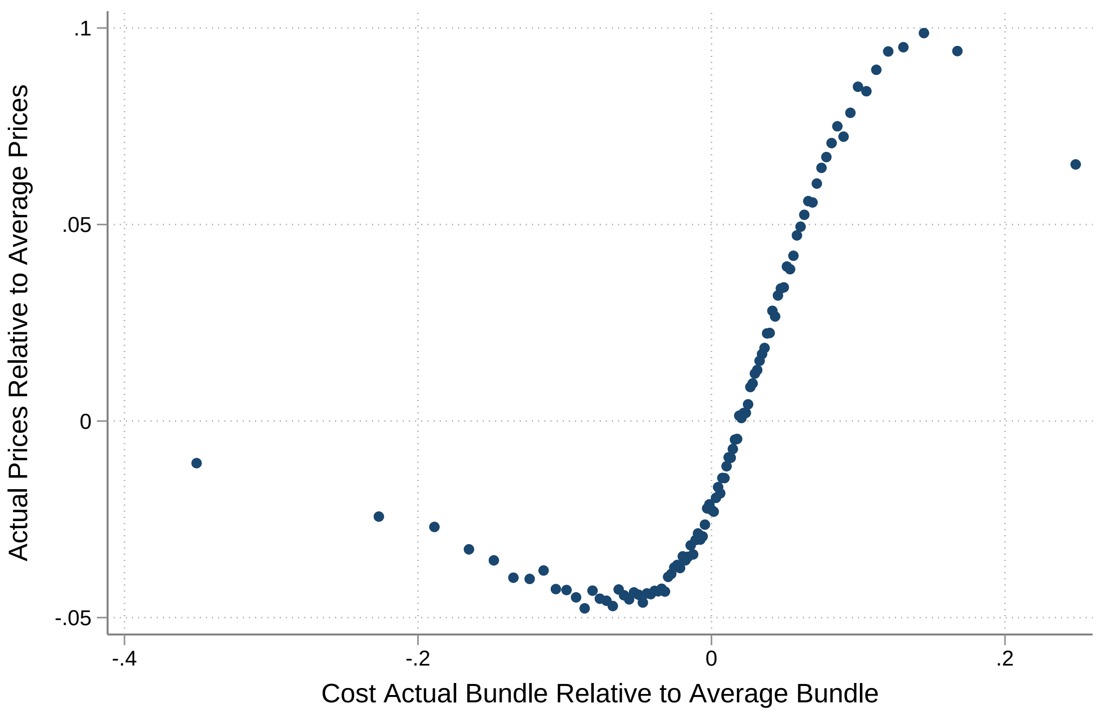
(**b**) Name-Brand Cage-Free Eggs: $3.69

(**c**) Store-Brand Organic Eggs: $3.99

(**d**) Name-Brand Pasture-Raised Eggs: $7.49

Notes: An illustration of the quality margin. This figure shows the prices of four different UPCs within the same product module at a single point in time. Quality adjustment means purchasing eggs that are cheaper or more expensive on a per-unit basis. Prices are actual prices at specific stores in Alameda County and *not* taken from the Kilts Nielsen Nielsen Panel.

Figure A.7: The Relationship Between Bundle Quality and Prices Paid



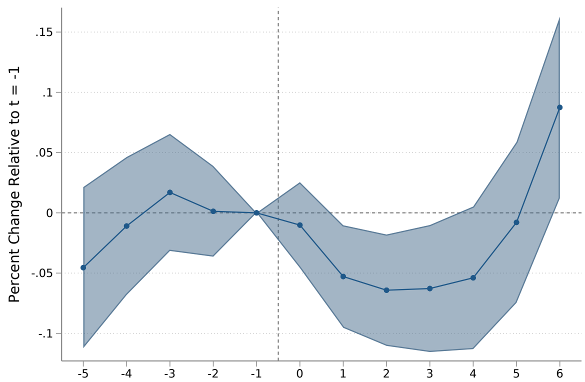
Notes: This figure presents a non-parametric binned scatter plot as proposed by Cattaneo et al. (2021). The effort and quality margins are as defined in Equations 7 and 10. The underlying data for this binned scatter plot is the baseline sample in Table 1.

Figure A.8: Decomposing the Post-Retirement Expenditure Drop

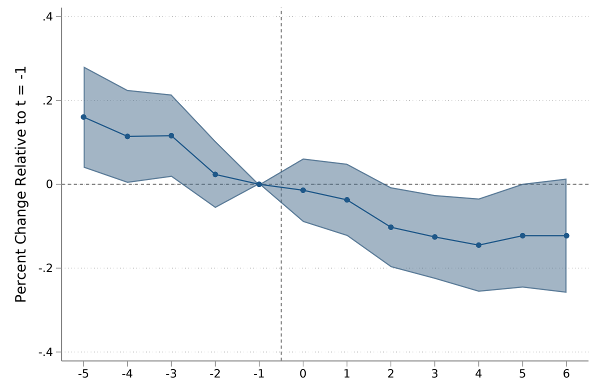
(**a**) Shopping Margins of Adjustment at Retirement

Notes: This figure presents the evolution of price effects, bulk effects, and quality adjustments estimated according to Equation 11 as well as uncorrected expenditure, expenditure at average prices and the average bulk allocation, and expenditure at average prices, the average bulk allocation, and average within-product module quality. The underlying data for each panel is the retirement sample of Table 1.

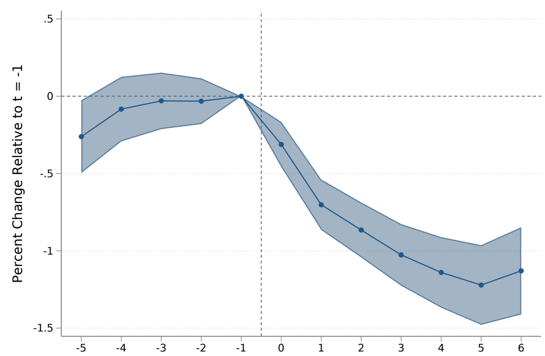
Figure A.9: Local Shopping Adjustments Around Retirement



(**a**) Prices Paid



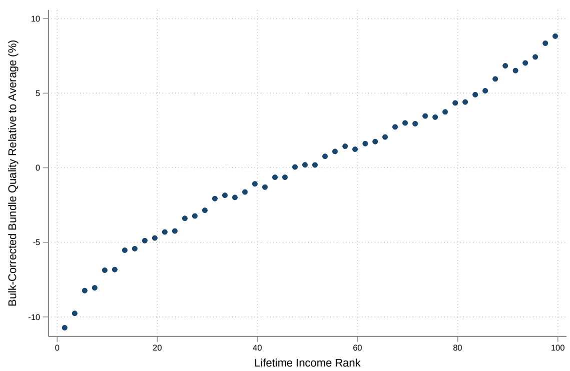
(**b**) Cost Relative to Average Bulk Allocation



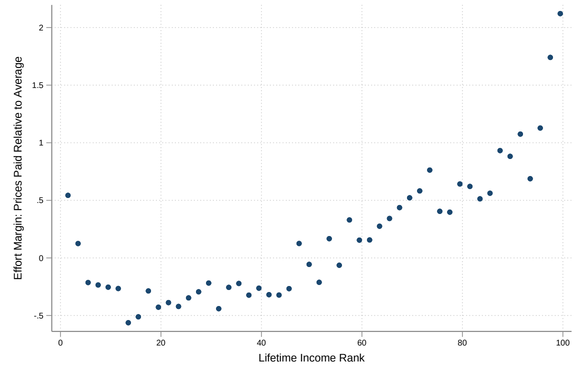
(**c**) Bundle Quality

Notes: This figure presents the evolution of price effects, bulk effects, and quality adjustments estimated according to Equation 11. All shopping margins of adjustment are defined in terms of deviations from local market averages. The underlying data for each panel is the retirement sample of Table 1. Standard errors are clustered at the household level.

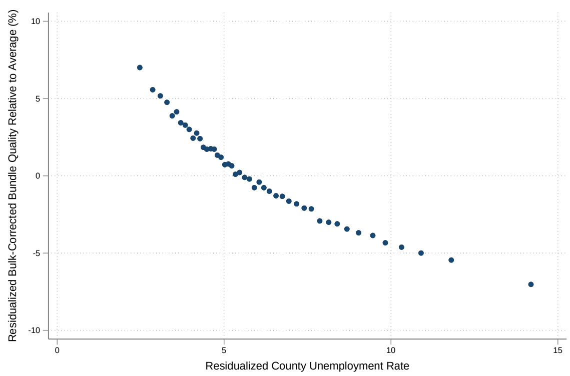
Figure A.10: Further Validations of Quality Measures



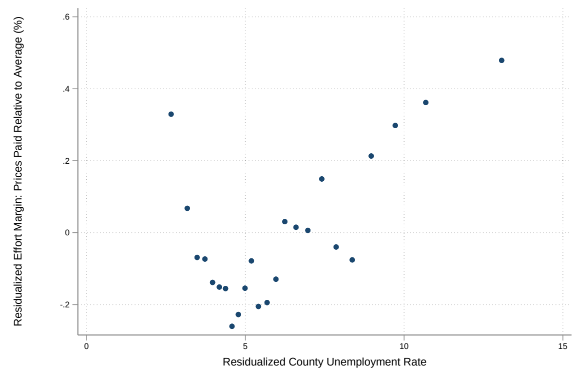
(**a**) Bundle Quality and Lifetime Income



(**b**) Prices Paid and Lifetime Income



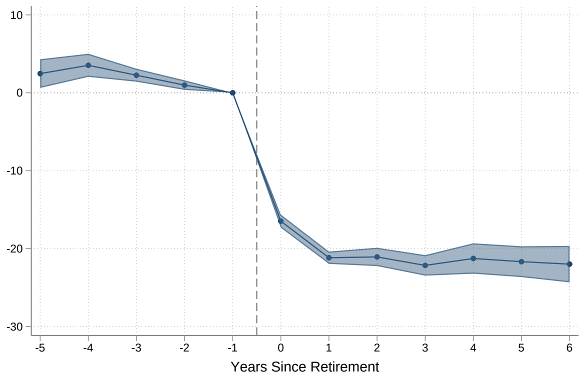
(**c**) Bundle Quality and Local Unemployment



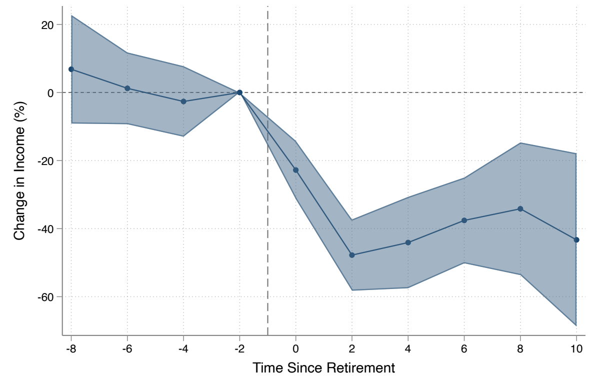
(**d**) Prices Paid and Local Unemployment

Notes: All figures present non-parametric binned scatter plots as proposed by Cattaneo et al. (2021) and control for household composition, household size, the age and presence of children, household age, a year fixed effect, and a scantrack market fixed effect. The effort and quality margins are as defined in Equations 7 and 10. The underlying data for Panels (a) and (b) is the set of households in the baseline sample who are in the data for at least three years; the data for Panels (c) through (d) is the baseline sample of Table 1.

Figure A.11: de Chaisemartin and D'Haultfœuille (2020) Estimates: Path of Income Around Retirement



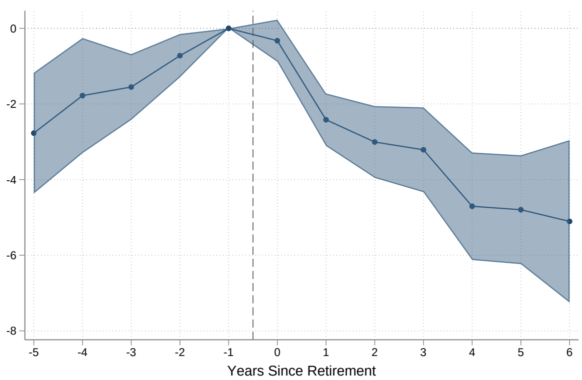
(a) Scanner Data



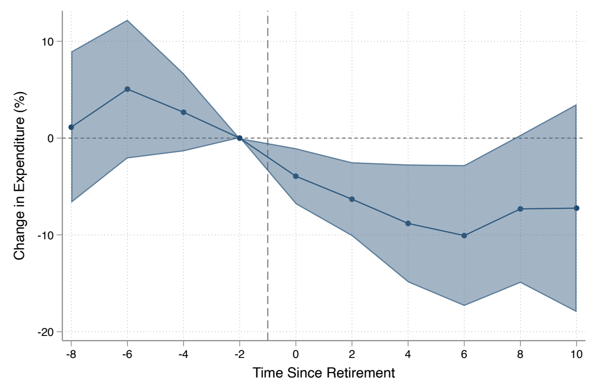
(b) Consumption Survey

Notes: This figure presents the evolution of income around retirement in the KNCP and HRS estimated with the event study estimator proposed by de Chaisemartin and D'Haultfœuille (2020). The underlying data for Panel (a) is the retirement sample of Table 1; the underlying data for Panel (b) is the retirement sample of Table 2. Bootstrap standard errors are clustered at the household level.

Figure A.12: de Chaisemartin and D'Haultfœuille (2020) Estimates: Expenditure Around Retirement



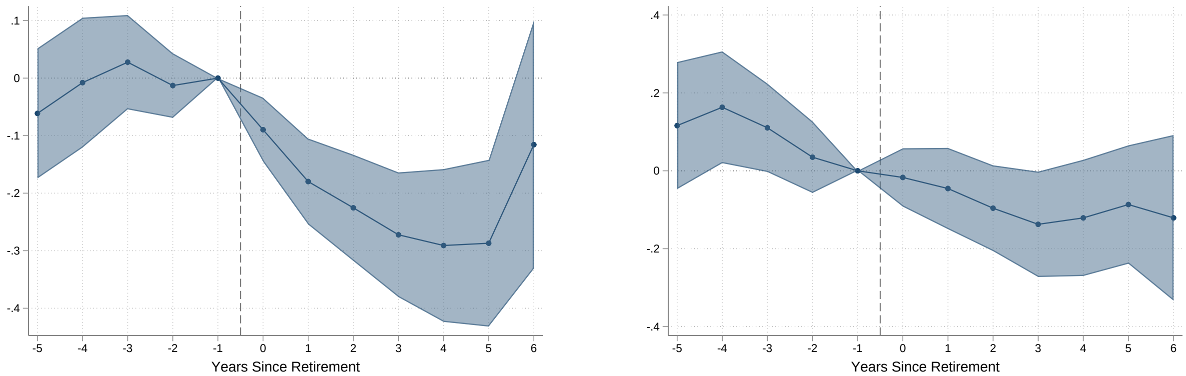
(a) Scanner Data



(b) Consumption Survey

Notes: This figure presents the evolution of expenditures around retirement in the KNCP and HRS estimated with the event study estimator proposed by de Chaisemartin and D'Haultfœuille (2020). The underlying data for Panel (a) is the retirement sample of Table 1; the underlying data for Panel (b) is the retirement sample of Table 2. Bootstrap standard errors are clustered at the household level.

Figure A.13: de Chaisemartin and D'Haultfœuille (2020) Estimates: Price and Bulk Savings Around Retirement

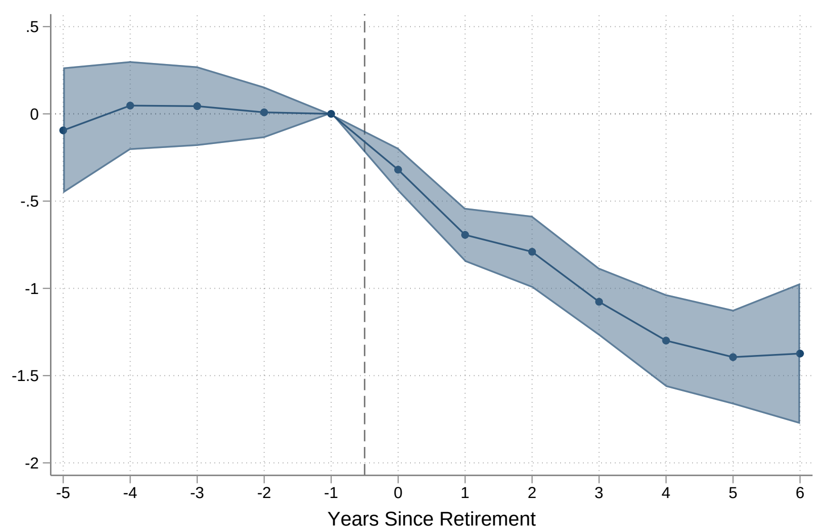


(**a**) Prices Paid

(**b**) Cost Relative to Average Bulk Allocation

Notes: This figure presents the evolution of price and bulk effects around retirement estimated with the event study estimator proposed by de Chaisemartin and D'Haultfœuille (2020). The underlying data for both panels is the retirement sample of Table 1. Bootstrap standard errors are clustered at the household level.

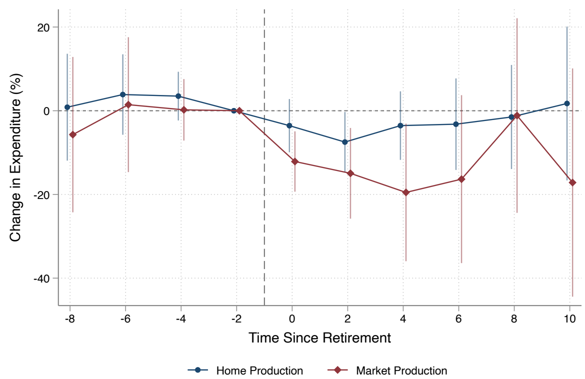
Figure A.14: de Chaisemartin and D'Haultfœuille (2020) Estimates: Consumption Quality Around Retirement



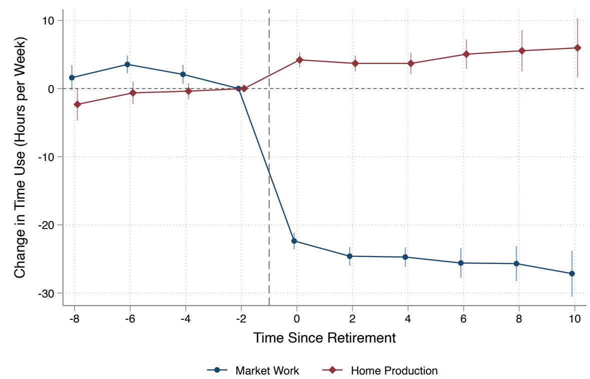
(**a**) Bundle Quality

Notes: This figure presents the evolution of quality adjustments around retirement estimated with the event study estimator proposed by de Chaisemartin and D'Haultfœuille (2020). The underlying data for both panels is the retirement sample of Table 1. Bootstrap standard errors are clustered at the household level.

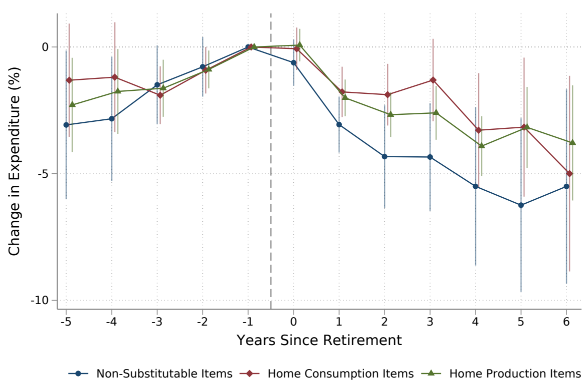
Figure A.15: de Chaisemartin and D'Haultfœuille (2020) Estimates: Home Production, Home Consumption and Expenditure Timing in Retirement



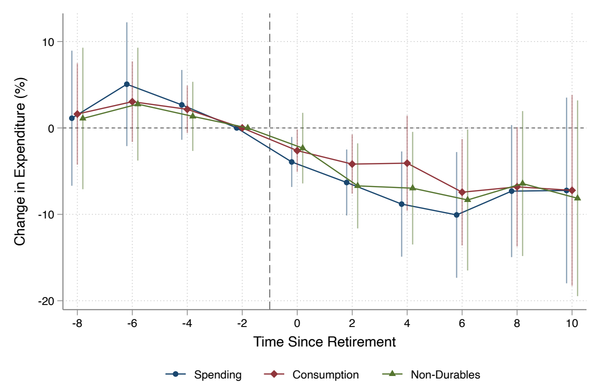
(a) Market and Home Production Expenditure



(b) Time Working vs. Home Production



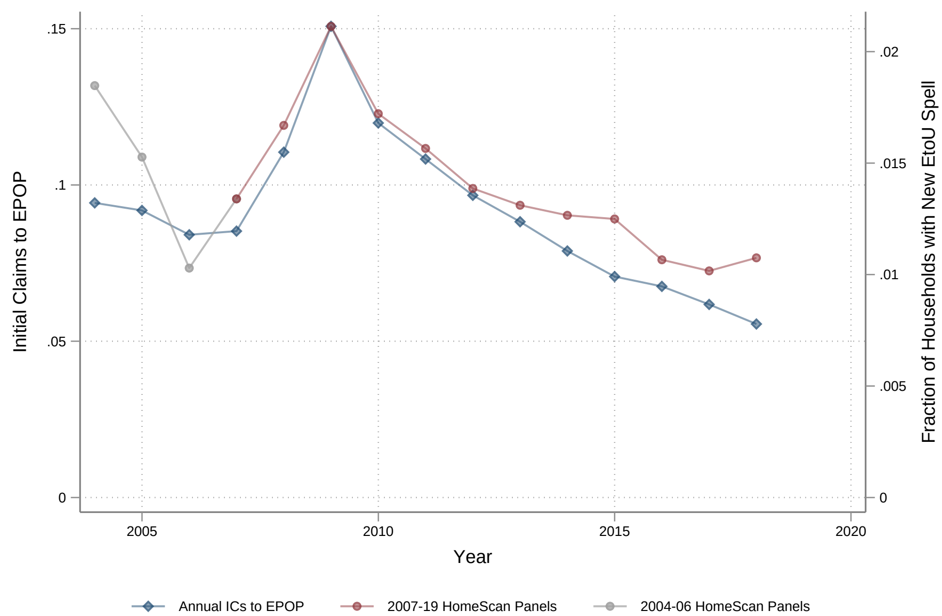
(c) Subsets of Spending by Substitutability



(d) Total vs. Consumption Expenditure

Notes: This figure presents the path of expenditure categories around retirement in the Kilts Nielsen Consumer Panel as well as expenditure categories and time use in the Health and Retirement Survey estimated according to Equation 11. Non-substitutable items are non-food grocery items, health and beauty items, and general merchandise. Home consumption items are prepared foods and deli items, food items that do not require preparation, and alcohol. Home production inputs are dry grocery items, packaged meat, fresh produce, frozen meat and vegetables, and dairy products. Home production items in the HRS are food at home, cleaning products, gardening products expenditures. Market production items are their immediate market good substitutes (food away from home, cleaning services, and gardening services). Home production time is time spent cleaning, doing gardening work, preparing meals and cleaning up and time spent shopping or running errands. The underlying data for Panels (a), (b), and (d) is the retirement sample of Table 2; the data for Panel (c) is the retirement sample of Table 1. Bootstrap standard errors are clustered at the household level

Figure A.16: Flows Into Non-employment in Nielsen and Aggregate Initial UI Claims



Notes: This figure compares flows into non-employment in the Kilts Nielsen Consumer Panel to the aggregate ratio of initial unemployment insurance claims to the employed population. In 2007, the Nielsen Nielsen Panel was expanded to include 50% more households and be a more accurate representation of national demographics.