Coming in at a Trickle: The Optimal Frequency of Public Benefit Payments *

Cameron LaPoint[†] Yale SOM

Shogo Sakabe[‡] Columbia

July 2023

Abstract

How governments should choose the frequency of payments has received little attention in the literature on the optimal design of benefits programs. We propose a simple model in which the government chooses the interval length between payments, subject to a tradeoff between administrative costs of providing more frequent benefits and welfare gains from mitigating recipients' consumption non-smoothing. Using a high-frequency retail dataset that links consumers to their purchase history, we apply the model to the Japanese National Pension System. Our evidence suggests suboptimal intra-cycle consumption patterns, with negligible retailer price discrimination. Our model calibrations support the worldwide prevalence of monthly payment schedules, even under extreme assumptions about preferences, and regardless of consumers' underlying behavioral frictions. For governments facing rapidly aging populations, our results imply lowering pension payment frequency may be a budget-preserving alternative to raising retirement age thresholds.

Keywords: optimal payment frequency, splurge goods, consumption smoothing, mental accounting, present bias, incidence, retail scanner data, pension reform

JEL classifications: D12, D91, G51, H21, H55, I38

*We would like to thank Michael Best, Judy Chevalier, James Choi, Harald Conrad (discussant), Anthony DeFusco, Wojciech Kopczuk, Kazuki Onji, Michaela Pagel, Daniel Reck, Matt Spiegel, Jón Steinsson, Neil Thakral, Takashi Unayama, Nathanael Vellekoop, Jialan Wang, Tsutomu Watanabe, David Weinstein, and participants in the Applied Micro Methods Colloquium at Columbia, Japan Empirical Economics Seminar at Osaka University, NBER Innovative Data in Household Finance, and Japan Economy Network (JEN) Conference at Columbia Business School for helpful advice and comments. We thank Louisa Liu for excellent research assistance. We are grateful to Magee Co., Ltd. (formerly, IDs Co., Ltd.) for providing the retail scanner data used in this project, and to the Research Institute of Economy, Trade and Industry (RIETI) for providing access to the JSTAR data. We also thank the Center on Japanese Economy and Business (CJEB) at Columbia Business School for supporting this project through the provision of dissertation fellowships to both authors. Additionally, Sakabe thanks the Nakajima Foundation for generous financial support. An earlier version of this paper circulated as CJEB Discussion Paper, No. 370.

[†]Yale School of Management. 165 Whitney Avenue, New Haven, CT 06511. Email: cameron.lapoint@yale.edu; Web: http://cameronlapoint.com

[‡]Department of Economics, Columbia University, 1022 International Affairs Building, 420 W 118th Street, New York, NY 10027. Email: s.sakabe@columbia.edu; Web: http://shogosakabe.github.io

1 Introduction

Most economies offer public benefits programs to provide regular transfer payments to their citizens. One important question that has been neglected in the literature on the optimal design of transfer programs is how the government should choose the frequency of disbursal to eligible households. Several studies have found that infrequently distributing benefits to people who live "paycheck to paycheck" may have adverse consequences. For instance, Dobkin & Puller (2007) document that recipients of Supplemental Security Income (SSI) drive spikes in drug-related hospital admissions at the beginning of the month. Similarly, Mastrobuoni & Weinberg (2009) find that Social Security recipients without savings consume fewer calories and switch to unhealthy foods near the end of the pay cycle.

In light of this evidence, we ask whether governments can improve household welfare by distributing transfer payments more frequently. To answer this question we propose a simple model that defines the optimal frequency of pension payments as a function of aggregate statistics about the benefits system and individual preferences. In our model, the government optimally chooses the length of the interval between benefit payments. This decision is subject to a tradeoff between the administrative cost of providing more frequent benefits and the welfare gain from reducing deviations from full consumption-smoothing behavior.

We apply our model to the Japanese National Pension System (JPS), which distributes bimonthly annuity payments that are a function of average monthly earnings while employed. Upon reaching retirement eligibility, contributors to the system can begin receiving pension payments every two months on the 15th of each disbursement month (February, April, June, etc.). However, if the scheduled delivery date falls on a Saturday, Sunday, or public holiday, payments are instead sent on the first previous non-holiday weekday. This timing rule, combined with annual variation in calendar weekdays, induces variation in the length of periods in between payments that is unrelated to pensioners' spending decisions.

We exploit this variation in the duration of payment cycles using a unique retail point-of-sale dataset from a Japanese marketing firm. Our sample includes the price and quantity of each good purchased in over 500 grocery stores across Japan over four years between 2011 and 2014. Transactions in our data are tied to loyalty point cards for which we observe a unique shopper ID and the birth month/year of the registered shopper. We are thus able to track individuals over time while observing exact prices paid for goods within each transaction at daily frequency.

Leveraging the high-frequency nature of the retail panel, we find that regular shoppers eligible for pension payments increase overall grocery expenditures by 10% within two days

of the scheduled delivery date. We also estimate a "duration elasticity," or the extent to which expenditure responses are more pronounced following longer intervals between payments in our sample time period. We use this duration elasticity and our estimates of the administrative costs of the JPS to calibrate our optimal payment frequency model. To estimate the shape of the administrative cost function, we exploit a reform to the Japanese pension system in 1988 that reduced the interval between payments from three to two months. The results from this natural experiment suggest the marginal cost to the government of increasing the payment frequency is negligible. We find annual costs rose by 4.3%, or by 0.14% per day the pay cycle decreased in length, for municipalities which contained JPS branch offices that bore the bulk of the administrative burden imposed by the reform, relative to municipalities without a branch office.

At the same time, we provide evidence that shoppers in our data can be described by a combination of quasi-hyperbolic discounters and myopic "payday liquid" consumers who exhibit a spike in consumption on payday but maintain an otherwise smooth consumption profile. For each type of agent, the implied marginal welfare loss is sufficiently non-trivial to justify an increase in the frequency of pension payments. We also consider alternative motivations for payday expenditures, including near-rationality and liquidity constraints. Following Kueng (2015, 2018), we use total expenditures over the pay cycle as a proxy for permanent income and document a slightly negative relationship between permanent income and payday expenditures, but only among shoppers with below-median permanent income.

We use our reduced-form estimates to parameterize the model and conclude that the optimal frequency of Japanese public pension payments is under one month, implying the government could improve welfare by increasing payment frequency. Moreover, even for a variety of underlying behavioral frictions and extreme assumptions about consumer preferences – such as a high intertemporal elasticity of substitution – or for programs with highly convex administrative cost functions with respect to interval length, our model rarely generates optimal frequencies exceeding one month. Incorporating preferences for consumption commitments (e.g. via durables) into our framework only slightly increases optimal pay cycle length. Hence, we offer a rationale for the worldwide prevalence of monthly public payment schedules.[1]

Using the Japanese old-age pension system as our setting helps avoid issues that have plagued the large literature on measuring the consumer response to anticipated payments.

[1]We collected information related to the administration of public benefits for 36 OECD countries from official government webpages. Of these, 25 disburse universal public pensions on monthly schedules, and eight distribute payments at even shorter intervals. Other forms of public benefits such as means-tested welfare programs are also commonly distributed on a monthly basis.

First, small living spaces in Japan make it difficult to buy in bulk or store groceries over long periods, so the transactions we observe more closely approximate "instantaneous consumption" rather than a savings mechanism. Second, universal health insurance coverage reduces the need of retirees to save their pension income for uncertain medical expenses. Third, as Stephens & Unayama (2011) document, over 80% of income for public pension recipients originates from these benefits. Since pensioners have to subsist for long intervals (two months) between which they receive little to no income, our application provides a particularly stark case where increasing the frequency of benefits might improve welfare.

We complement papers examining the frequency of pay by offering a sufficient statistics approach to computing optimal pay frequency from a regulator's point of view. Maloch & Weaver (1969) were among the first to note in their study of the spending patterns of 26 Ohio families that households paid weekly find it easier to save. Shapiro (2005) posited that policymakers might improve welfare by increasing the frequency of benefit payments. Parsons & Van Wesep (2013) and Baugh & Correia (2022) build on this notion in studying how workers' pay cycle duration influences consumption and borrowing patterns. Dobkin & Puller (2007) argue that recipients of SSI drive spikes in drug-related hospital admissions at the beginning of the month, while others show that crime increases (Foley 2011) and mortality falls (Evans & Moore 2012) towards the end of monthly welfare cycles. The link between physical harm and non-smooth consumption paths at higher frequencies suggests estimates of the utility loss from permanent income hypothesis deviations, as in Cochrane (1989) and Browning & Crossley (2001), may be severely underestimated.

Our empirical application tests whether the magnitude of deviations from the permanent income hypothesis varies with the timing of *regular* income sources – that is, recurring payments, such as pension payments, that are delivered more than once per year. Stephens & Unayama (2011) look at the Japanese pension system during an earlier period in the 1980s when pension payments were distributed quarterly. Consistent with our results, they provide suggestive evidence that consumption growth in the month of check receipt is lower after the reform which shifted from a quarterly to bimonthly payment schedule.[2] A more recent strand of household finance research uses high-frequency data from fintech and bank accounts to document consumer responses to various income sources. Gelman et al. (2014) find that total spending rises 70% above the daily average on the day a regular paycheck arrives. Olafsson

[2]Less recent studies use quarterly or monthly panel data and examine payments that occur at the cutoff separating two discrete time observations (e.g. Browning & Collado 2001; Stephens 2003, 2006). For example, with monthly panel data, estimates of consumption growth between June and July due to receipt of a bonus at the end of June are attenuated if households spend a portion of the bonus in June and the other portion in July. We overcome this issue by using high-frequency scanner data that allows us to observe expenditure responses at a precise number of days between scheduled payment dates.

& Pagel (2018) find similar results in Iceland and argue that consumers act as if they have a license to spend at the beginning of a new pay cycle. Baker (2018) shows that liquidity constraints can explain heterogeneity in the responsiveness of consumption to income shocks. Using a similar empirical setting to ours, Baugh & Wang (2021) exploit within-household variation in the length of Social Security pay cycles to show that households are more likely to experience financial shortfalls during longer pay cycles. While those authors argue that inattention to changes in the length of the pay cycle can explain their results, we provide evidence from Google search data that consumers in our setting are highly attentive to even small changes to usual disbursement dates.

Our analysis of expenditure patterns of the elderly also adds nuance to the literature on the retirement consumption puzzle. Bernheim, Skinner, & Weinberg (2001) show that total expenditures drop at retirement for consumers in all but the highest income and wealth quartiles. Aguiar & Hurst (2005) contend that this drop does not imply sub-optimal saving for retirement, as at-home food production offsets the decline in food expenditures.[3] After accounting for heterogeneity in retiree cohorts à la Sun & Abraham (2021), we find shoppers increase their consumption of perishables in the months after crossing retirement age thresholds. This indicates savings stocks are likely to be high among benefit recipients.

While we cannot directly test for liquidity constraints, our results imply at least part of the spending response to payday is due to behavioral factors rather than short-term liquidity constraints. We document similar payday responses for shoppers who visit stores at different average shopping frequencies. Coibion, Gorodnichenko, & Koustas (2021) show that the ability of upper-income households to buy in bulk helps explain increasing U.S. expenditure inequality. Based on the insight of Bils & Klenow (2001) that quality Engel curves slope upward, we rank stores based on average prices of regularly-transacted goods and match shoppers to their most-visited store as a tag for income and find nearly identical payday spending responses when we split shoppers along that dimension.

Whether retailers capture some of the incidence of public benefit payments by engaging in price discrimination against eligible consumers is an open question. This question is of direct relevance to the welfare calculations we conduct here, since increasing the frequency of payments may also inhibit retailers' ability to engage in price discrimination if menu costs are

[3]Other papers in this literature include Battistin et al. (2009), who estimate a drop in consumption due to male retirement in Italy. Stephens & Unayama (2012) find little evidence of a drop in consumption among Japanese retirees, but show that this is primarily due to working households receiving large lump-sum retirement bonuses from their employer. In contrast to the majority of studies which rely on expenditure survey data, Agarwal, Pan, & Qian (2015) and Olafsson & Pagel (2020) analyze transactions-level data linked to bank accounts in Singapore and Iceland, respectively. We instead use retail point-of-sale data that allows us to separate prices and quantities, thereby isolating consumption from expenditures.

sufficiently large. Warner & Barsky (1995) collect daily price data for retail stores between November and February and document that markdowns tend to occur during more intensive shopping periods. This finding is echoed by MacDonald (2000) and Chevalier, Kashyap, & Rossi (2003), who argue that average prices fall during seasonal demand peaks due to coincident declines in retail margins. Nevo & Hatzitaskos (2006) instead offer consumer substitution towards an increased supply of cheaper products as an explanation. In contrast, Hastings & Washington (2010) show that the price index for the basket of food goods consumed by SNAP-recipient households in Nevada falls by 3% over the month. Goldin, Homonoff, & Meckel (2022) extend the analysis to 48 states and a large number of stores and find no evidence of a retailer pricing response on SNAP delivery dates.

Overall, even though our model can easily incorporate strategic retailer responses, altering the pay cycle length is unlikely to affect retailers' ability to capture the incidence of benefit payments. We find limited evidence that retailers capture the incidence of pension payments by price discriminating against elderly customers on or around scheduled delivery dates. Average prices paid within a store increase by 1.5% on payday, but the number of unique goods purchased increases by 5.9%. Using a price index that isolates a change in average prices due to consumer substitution from the retailer's pricing response, we demonstrate that the observed spike in store-level prices is almost entirely due to consumers substituting towards an expanded expenditure basket which includes higher quality goods, rather than retailers raising prices to capture the incidence of pension payments.

This splurge behavior on payday reinforces the theory, introduced in Chevalier & Kashyap (2019), of consumer "type switching" from bargain hunters to brand loyalists during periods of peak demand. The substitution and variety effects we uncover are also consistent with evidence on the non-fungibility of SNAP benefits (Hastings & Shapiro 2018) and tax refunds (Baugh, Ben-David, & Parker 2021), and with the literature on mental accounting (Thaler 1999; Farhi & Gabaix 2020) which argues individuals earmark income sources for specific spending categories. Our finding of within-retailer substitution and variety effects complements the analysis in Baker, Baugh, & Kueng (2021), who show households switch to higher quality retailers when their income increases. In documenting that consumer purchase decisions are important drivers of store-level inflation during peak demand periods, our results on pay frequency suggest the staggered rollout of payments can mitigate inflationary pressures induced by fiscal stimulus policy (Sahm, Shapiro, & Slemrod 2012, 2015).

Finally, pension reforms proposing to raise the normal retirement age have taken center stage due to population aging and the resulting financial strain on social security systems in developed countries (Kolsrud et al. 2023). While our framework applies generally to

all public benefit programs, our results imply lowering pension payment frequency may be a budget-preserving – but more politically feasible – alternative to raising the normal retirement age. At the same time, we show increasing the normal retirement age puts upward pressure on the optimal frequency, generating extra cost savings from reduced operating expenditures in addition to the oft-cited benefit of increased income tax revenues from maintaining a larger working population (Gruber & Wise 1999).

The remainder of the paper is organized as follows: Section 2 presents our optimal payment frequency model. Section 3 provides background on the Japanese pension system and describes our data. Section 4 presents our main empirical results on the expenditure response to payment receipt. Section 5 assesses possible retailer pricing responses and shows consumer substitution towards higher quality goods. Section 6 discusses our calibration of the model to determine the optimal payment frequency. Section 7 concludes.

2 A Simple Optimal Payment Frequency Model

In this section, we outline the modeling framework for the government's choice of the optimal interval between regularly occurring public benefit payments. While we present the model in the context of a public pension system, the model can be applied to a variety of policy contexts where the government disburses payments at regular, anticipated, intervals.

2.1 Basic Framework

Consider periods of time t that occur within the interval $[0, T]$, where each unit of time is measured in days. A fraction p of people in the economy are pensioners who receive a flat pension benefit every T days equal to $b(T) = \overline{B} \cdot T$. The other $1-p$ fraction of people in the economy are workers, who instead of receiving the pension benefit, earn an arbitrary wage $w(t)$ and pay a lump-sum tax of $\tau(b)$ that is used to finance the pension system.[4]

In addition to the direct cost of $p \cdot b(T)$ of delivering benefits to retirees, the government faces an administrative cost function $\mu(T)$.[5] While the shape of $\mu(T)$ will ultimately depend on the particular application, for now we assume the cost function is both strictly increasing

[4]The portion of JPS benefits that are not contingent on employment are financed through monthly lump-sum payments. The value of the monthly payment in 2018 was 16,340 JPY ($\approx$ \$143).

[5]These administrative costs to implementing the pension system may include costs associated with: authorizing benefits, delivering benefits (either electronically or through post), redeeming and reconciling benefits, investigating and prosecuting fraud, managing pension funds.

in the frequency of payments and weakly convex, so $\mu'(T) < 0$, $\mu''(T) \geq 0$. We discuss our strategy for estimating the administrative cost function in Section 6.

At each point in time, assume the government must follow a balanced budget rule, so that for each $t \in [0, T]$ the following must hold:

$$(1-p) \cdot \tau(b) = p \cdot b(T) + \mu(T) \implies \tau(b) = \frac{p \cdot \overline{B} \cdot T + \mu(T)}{1-p} \tag{2.1}$$

Working households have instantaneous utility of consumption $u(C(t))$. These non-pensioners face a standard consumption-saving problem of

$$\max_{\{C(t)\}} \int_0^T u(C(t))dt \text{ s.t. } C(t) = S(t) + w(t) - \frac{\tau(b)}{T} \tag{2.2}$$

where for simplicity we normalize $r = 0$. In the absence of liquidity constraints that would place restrictions on the asset position $S(t)$, the non-pensioner's optimal consumption path is completely flat over the interval: $C(t) = C^*, \forall t \in [0, T]$.

Let $u^r(C(t))$ denote the instantaneous utility of retirees. Pensioners are given $b(T)$ to spend over the payment cycle. The optimal consumption path of retirees is therefore also constant on the interval $[0, T]$. However, we suppose that, following Shapiro (2005), the pensioner instead chooses a consumption path given by

$$C(t) = \exp\Big(\theta - f(t)\Big) \tag{2.3}$$

where $f(t)$ is a potentially non-monotonic function representing the deviation of the chosen consumption path at each time t from a constant value. Because the source of this deviation will again depend on the particular application, in the general setup we remain agnostic as to the underlying behavioral phenomenon driving the household away from the consumption smoothing benchmark. The only restriction we impose on $f(t)$ here is that $f(0) = 0$. Since the only income pensioners receive within the payment cycle is $b(T)$, the budget condition pins down the value of the constant θ:

$$\int_0^T \exp\Big(\theta - f(t)\Big) = b(T) \implies \exp(\theta) = \frac{b(T) \cdot f'(0) f'(T)}{f'(T) - f'(0) \cdot \exp(-f(T))} \tag{2.4}$$

To determine the potential welfare gain that the government might achieve by changing payment frequency, we can ask the share $(1 - \lambda)$ of its benefit the pensioner household would be willing to give up to achieve the optimal constant consumption profile. This can

be obtained as the fraction λ that solves the following equation:

$$\int_0^T u^r\Big(\exp\Big(\theta - f(t)\Big)\Big)dt = \int_0^T u^r(\lambda\overline{B})dt \tag{2.5}$$

Shapiro (2005) solves this expression for λ in the special case where $f(t) = \nu{\cdot}t$ and $u^r(C(t)) = \frac{C(t)^{1-\rho}}{1-\rho}$, so that the welfare loss is independent of the benefit level. Our aim here is more general: we wish to find the optimal frequency T that minimizes the welfare loss, given the total costs of funding the pension system.

Suppose there exists a constant γ that represents the cost $\tau(b)$ of a dollar spent by the government in the same units as total utility of the non-pensioners over the pay cycle. Such a constant corresponds to the marginal cost of funds (MCF), and under a lump-sum tax imposed only on the non-pensioners, $\gamma = 1$.[6] Given a constant MCF, the optimal frequency problem a utilitarian government faces is given by

$$\min_T \Big\{ -p\cdot\lambda(T) + \gamma\cdot\Big(p\cdot b(T) + \mu(T)\Big)\Big\} \tag{2.6}$$

where $\lambda(T)$ is the solution to (2.5). For any strictly concave $u^r(\cdot)$, the welfare loss $1-\lambda$ will be strictly convex in T. Therefore, in such a case the FOC of this problem is necessary and sufficient for a solution:

$$\frac{p\cdot\lambda'(T^*)}{\gamma} = \mu'(T^*) + p\cdot\overline{B} \tag{2.7}$$

We have now obtained T^* as an implicit function of observables. These include the daily average benefit amount among claimants over the pay period $\overline{B} \equiv b(T)/T$ and the fraction p of the population who are claimants. In most cases, $b(T) = \overline{B}\cdot T$, is linear since the total benefit amount disbursed to claimants does not depend on the pay cycle length. In what follows we focus our attention on this scenario.

Next we turn to two special cases of the model which are motivated by observations in the literature on expenditure responses to income receipt. One case features benefit recipients who are quasi-hyperbolic discounters, and another assumes recipients behave as if they have a license to spend on payday but maintain an otherwise smooth consumption profile.

[6] More generally, for a distortionary tax τ, the MCF can be computed as $\gamma = -\frac{\partial\mathcal{U}^*/\partial\tau}{\partial R^*/\partial\tau}$ where $\mathcal{U}^*$ is the total optimized level of utility for the non-pensioners, and R^* is revenue collected at the optimum.

2.2 Quasi-hyperbolic consumers

We describe a special case of our optimal frequency model where pensioners are quasi-hyperbolic (hereafter, QH) discounters à la Laibson (1997), and hence the deviation function $f(t)$ is linear in time. The analysis here builds on the special cases examined in Shapiro (2005) and Mastrobuoni & Weinberg (2009), who posit that consumer preferences take the form:

$$u(c_0) + \beta \sum_{t=1}^{T} \delta^t u(c_t) \tag{2.8}$$

where β is the QH discount factor, and δ is the standard daily exponential discount factor. Both papers focus on the case where $\delta = 1$, $\beta < 1$, and the felicity function $u(\cdot)$ is isoelastic.[7] Under these conditions, one can show that the decrease in log consumption over time is:

$$\frac{\partial \log(c_t)}{\partial t} = \frac{1}{\rho} \cdot \log \beta - \frac{1}{T-t+1} + \frac{1}{T-t+\beta^{-1/\rho}} < 0 \tag{2.9}$$

where ρ is the inverse intertemporal elasticity of substitution (IES). For values of β close to one, the time path for consumption given by (2.9) is approximately linear over the pay cycle.[8] Thus following Shapiro (2005), in the setup of our optimal frequency model we capture QH discounting behavior for the case where $f(t) = \nu \cdot t$ and ν is the constant daily rate of decline in consumption over the pay cycle. The linear rate ν corresponds to some combination of β and ρ that rationalizes an observed total percentage decline in consumption over the pay cycle. For instance, Shapiro (2005) finds that consumption declines by 0.4% daily ($\nu = 0.004$) for the average household participating in Maryland's food stamps program. This is comparable to the $\nu = 0.006$ estimate implied by Huffman & Barenstein's (2005) analysis of monthly pay cycles in U.K. expenditure survey data.

To illustrate, for log utility ($\rho = 1$), the compensating variation term from (2.5) is

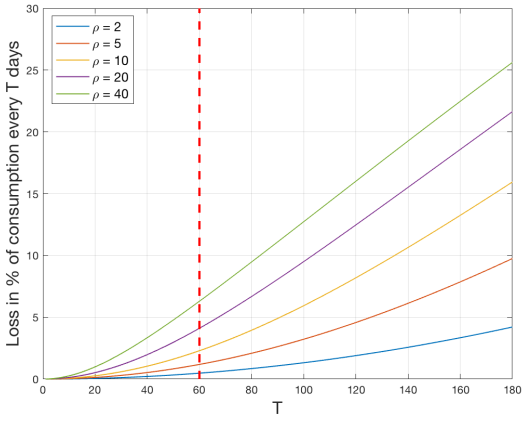
$$\lambda(T) = \frac{1}{\overline{B}} \cdot \exp\left(\theta - \nu \cdot T/2\right) \tag{2.10}$$

[7] The case of $\delta = 1$ is a reasonable approximation to the daily exponential discount factor for analyzing consumption decisions over a relatively short interval between payments.

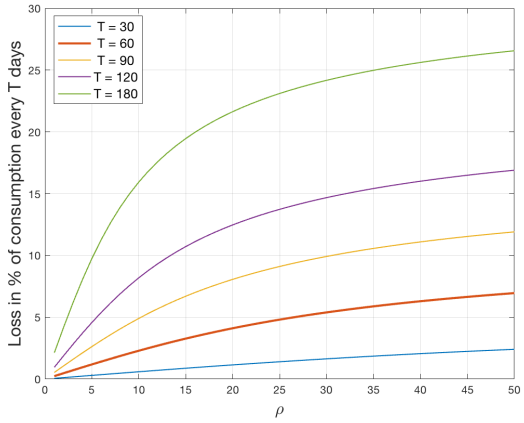
[8] Exponential discounters with $(\beta, \delta) = (1, 1)$ have a flat consumption profile, and for $\beta = 1$, $\delta < 1$ the exponential discounters have a linear decline in consumption over the cycle with slope $\log(\delta)/\rho$.

FIGURE 1. Welfare Loss under Quasi-Hyperbolic Discounting

A. By Interval Length



B. By Inverse IES



Notes: The figure plots the welfare loss implied by quasi-hyperbolic discounters as a function of the government's choice of interval length (left), or as a function of the inverse intertemporal elasticity of substitution (right), under isoelastic utility. Full derivations and closed-form expressions for the welfare loss can be found in Appendix A.2.

and the FOC to the government's problem is given by

$$-\frac{p \cdot \nu}{2\overline{B}} \cdot c_0 \cdot \exp\Big(-\nu \cdot T/2 \Big) = \mu'(T) + p \cdot \overline{B} \tag{2.11}$$

We provide a more general expression for the welfare loss under isoelastic utility with $\rho \neq 1$ in Appendix A.

In Section 6 we numerically solve for the welfare loss and the optimal payment frequency from a calibrated version of our model with pensioners as QH discounters. For empirically valid combinations of ν and ρ, the welfare loss from non-smoothing is relatively small for Japanese pension payments. For example, if the average daily consumption decline over the pay cycle is 0.4%, then for $\rho = 10$, the welfare loss with a bimonthly pension system ($T = 60$) is 2.3% of consumption; in the log utility case the loss is only 0.24% of consumption.[9]

Figure 1 summarizes how the welfare loss varies with the interval T for different assumed values of the inverse IES. We note two features about the welfare loss under QH discounting that the figure illustrates. First, the welfare loss is increasing in the government's choice of the interval length T. Intuitively, for longer pay cycles the integral between the optimal smooth path and the path under QH discounting will be greater, as log deviations from the smoothed

[9]Most non-experimental estimates for the inverse IES fall in the range $\rho \in [1, 3]$. However, Best at al. (2020) use notches in mortgage interest rate schedules in the U.K. to argue that $\rho = 10$ is plausible.

level of consumption grow linearly over time. Second, the welfare loss is increasing in the inverse IES. A higher value for ρ means consumption is less substitutable between periods, so an individual is willing to pay more *ex ante* to get closer to the smooth consumption path.

2.3 Payday liquid consumers

We now examine a version of our model where we assume pensioners exhibit "payday liquidity," or mental accounting. Several recent studies exploiting high-frequency transaction data have documented that consumers spend considerably more around the day they receive a regular paycheck or benefit payment from the government. For instance, Gelman et al. (2014) analyze 60 million transactions in the U.S. and find that, on average, total spending rises 70% above the daily average on the day a regular payment arrives. Olafsson & Pagel (2018) report similar results from a financial planning app in Iceland, with the poorest tercile of households in their sample spending 70% more on a payday, and the richest tercile spending 40% more on paydays. Both papers show that the spike in consumption on paydays cannot be fully explained by measures of liquidity. Households instead act as if they have payday liquidity, or a license to spend at the beginning of a new pay cycle.

In our framework, we consider a case where pensioners heuristically spend more at $t = 0$ when they receive their payment, but their consumption path is otherwise smooth over the rest of the pay cycle. Due to the discontinuous nature of the consumption path in this setting, in this subsection we proceed in discrete rather than continuous time for ease of exposition.[10] We assume pensioners consume over the time interval $[0, T-1]$ as follows:

$$C_t = \begin{cases} (1+x)\cdot \bar{c} & \text{if } t = 0 \\ \bar{c} & \text{if } t \in [1, T-1] \end{cases} \tag{2.12}$$

As before, pensioners receive $\overline{B} \cdot T$ to spend over the time period $[0, T-1]$, so that the budget constraint pins down the value of $\bar{c}$.

$$\begin{aligned} &\sum_{t=0}^{T-1} C_t = (1+x)\cdot \bar{c} + \sum_{t=1}^{T-1} \bar{c} = (T+x)\cdot \bar{c} = \overline{B} \cdot T \\ &\implies \bar{c} = \begin{cases} \frac{\overline{B}\cdot T}{T+x} & \text{if } T > 1 \\ \overline{B} & \text{if } T = 1 \end{cases} \end{aligned} \tag{2.13}$$

[10] It is possible to directly map the results in discrete to continuous time using the concept of Cauchy principal values to evaluate integrals of functions that are discontinuous over the time interval.

$\bar{c}$ represents the constant value of consumption in all days within the pay cycle besides the payday where consumption spikes. The value $x = c_0/\bar{c} - 1$ represents the magnitude by which consumption spikes on payday.[11]

Many utility functions can produce the consumption path in (2.13) as the solution to utility maximization. For instance, a log felicity function with a $(1+x)$ weight on utility in $t = 0$ will generate a spike on payday of $x\%$ relative to $\bar{c}$.[12] However, if C_t were consistent with utility maximization, then there would be no welfare loss from non-smoothing, and our work here would be done. We consider a more interesting case of myopic consumers who heuristically spend more on payday, even though this is inconsistent with utility maximization.

The discrete time analog of the compensating variation equation in (2.5) is:

$$\sum_{t=0}^{T-1} u(C_t) = T \cdot u(\lambda \overline{B}) \tag{2.14}$$

To facilitate comparison to other cases of our model, we continue to assume an isoelastic felicity function with inverse IES ρ. The expression for the welfare loss is then:

$$1 - \lambda(T) = \begin{cases} 1 - \frac{\bar{c}}{\overline{B}} \cdot (1+x)^{1/T} & \text{if } \rho = 1 \\ 1 - \frac{\bar{c}}{\overline{B} \cdot T^{1/(1-\rho)}} \left[(1+x)^{1-\rho} + (T-1) \right]^{\frac{1}{1-\rho}} & \text{if } \rho \neq 1 \end{cases} \tag{2.15}$$

Figure 2 plots the welfare loss formula in (2.15) for the isoelastic payday liquid consumer as function of T for different values of ρ. In this figure we assume a conservative estimate for the spike in payday consumption ($x = 0.1$) that we obtain in Section 4.2 for perishable food consumption. We normalize the welfare cost to be 0 when there is only one period. The welfare loss with payday liquid pensioners is declining in the government's choice of T (Panel A). Because the welfare loss is entirely concentrated in the initial drop in consumption between periods $t = 0$ and $t = 1$, increasing the length of the pay cycle merely subdivides this loss over a larger total amount of consumption. Hence, for any non-negative cost function $\mu(T)$, the optimal payment frequency is a corner solution. Even if the government incurs zero costs to distributing payments more frequently, it could minimize the welfare loss by choosing the largest T^* such that pensioners can subsist ($\bar{c} > 0$).

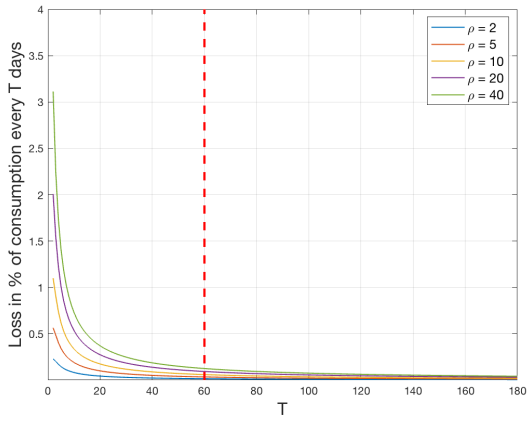
[11] While in practice consumption is not literally flat over time intervals with no paydays, it is roughly flat net of a full set of calendar effects (day-of-the-week, week-of-the-month, month-year). We show results to this effect later in the paper. In that sense, the consumption path we model here is that of average daily consumption, or an ergodic process of consumption.

[12] More generally, an isoelastic felicity function with a $(1+x)$ weight on utility in $t = 0$ will generate a spike on payday of $[(1+x)^{1/\gamma} - 1]$ % relative to $\bar{c}$.

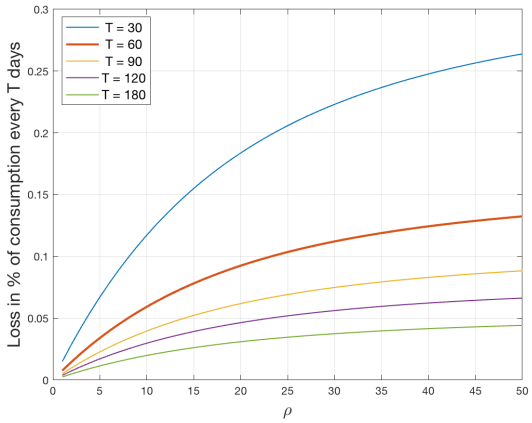
FIGURE 2. Welfare Loss under Payday Liquidity

A. Standard Payday Liquidity (Constant x)

By Interval Length

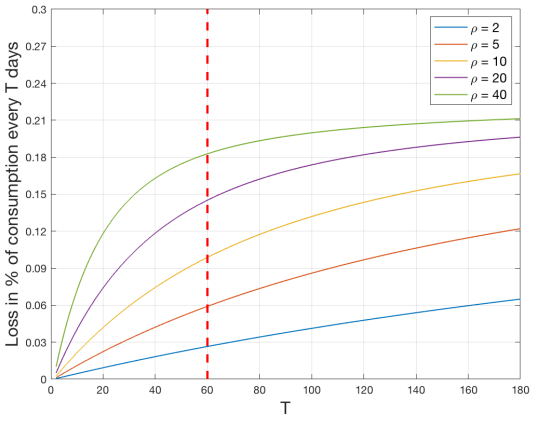


By Inverse IES

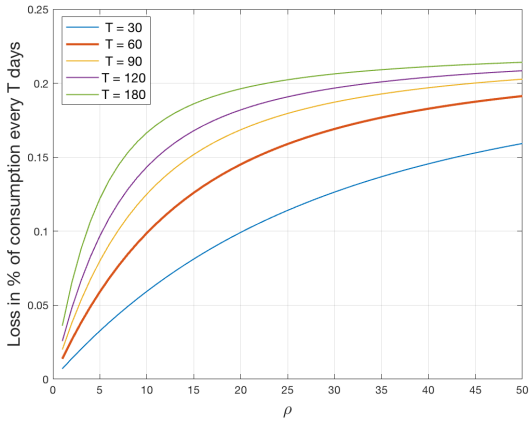


B. Payday Liquidity with Pent-up Demand (x linear in T)

By Interval Length



By Inverse IES



Notes: Panel A of the figure plots the welfare loss in equation (2.15) implied by payday liquidity behavior as a function of the government's choice of interval length, or as a function of the inverse intertemporal elasticity of substitution, under isoelastic utility. Panel B plots the same object when we allow the spike magnitude $x(T)$ to vary with interval length, which generates pent-up demand at the beginning of pay cycles. See Appendix A.3 for detailed derivations.

So far we have assumed that the spike in consumption on payday is independent of the interval between payments. If instead $x'(T) > 0$ this would capture the notion that consumers splurge more on a payday when more time has passed since they last received a payment. In other words, there may be more "pent-up demand" when pay cycles are longer. One can show that in this scenario decreasing T can improve welfare if the loss from the increase in the spike magnitude associated with an increase in T exceeds the welfare gain from subdivision as T increases. With log utility that condition reduces to the following:

$$\underbrace{\Big(1 + x(T)\Big) \cdot \log\Big(1 + x(T)\Big)}_{\text{loss from spike magnitude as } T\uparrow} \quad > \quad \underbrace{T \cdot x'(T)}_{\text{gain from subdivision as } T\uparrow} \tag{2.16}$$

In our empirical setting, we estimate a slope of $x'(T) = 0.0013$ by exploiting exogenous calendar variation in pension delivery dates, so the spike on payday increases by 0.13 percentage points for each additional day in the pay cycle. For this value of $x'(T)$ the condition in equation (2.16) holds for any $\rho > 0$. Panel B of Figure 2 shows the welfare loss as a function of T under this calibration. In contrast to the case where the spike was assumed to be a constant, the welfare loss is now increasing and concave in interval length.

The welfare loss for payday liquid consumers is much smaller than that for QH discounters when $T = 60$ as in the current Japanese Pension System. This is true regardless of whether the spike is a function of T or a constant. For example, at $T = 60$, for $\rho = 10$ and $x = 0.1$ the welfare loss for payday liquid consumers is 0.06%, compared to 2.29% of consumption for QH discounters. When we allow the spike on payday to depend on the interval length, the loss increases to 0.1% of consumption. The small magnitude of the loss is unsurprising given that the welfare loss for payday liquid consumers is entirely concentrated in the initial drop in consumption. Yet, what matters for the optimal frequency is the *marginal* welfare loss. We show in Section 6 that the marginal welfare loss is sufficiently non-trivial in each of these cases to justify an increase in the frequency of payments in the JPS.

To summarize this section, we have characterized two special cases of our optimal payment frequency model that feature prominently in the literature: QH discounting and payday liquidity. We present full derivations and model extensions in Appendix A, and assess the empirical relevance of these two special cases and alternative theories in Section 4.4. In Appendix A.4, we extend our optimal frequency results to scenarios where consumers are "sophisticated," meaning they internalize their tendancy to over-spend on payday.

3 Data and Background

In this section we provide background on the Japanese National Pension System and describe the retail panel and municipal-level data used to construct administrative cost measures.

3.1 Data

Retail panel data. We use retail scanner data provided by Magee Co., Ltd. to measure how expenditures and retail prices respond to shoppers receiving payments from the Japanese National Pension System (JPS).[13] The data record sales receipts from 1,120 grocery stores of 19 grocery chains across Japan from April 2011 to October 2014. Each sales receipt consists of the purchase date, a consumer-identifier code, a store-identifier code, and prices and quantities of barcode-level purchases. An unique characteristic of this dataset is that it includes consumer identifiers assigned at the time consumers sign up for a shopper loyalty program. This shopper ID has information about each consumer's birth year/month and gender. If a consumer uses their membership card when shopping, their identifier code is recorded on the sales receipt, which enables us to track their purchasing history over time.

Although the data include all transaction records for each store during the sample period, we focus on transactions involving consumers who own and regularly used their membership cards, as this allows us to assign pension eligibility status. We define regular shoppers as those who use their point cards at least four times per month (eight times per pay cycle) starting from the first month they appear in the data. This leaves us with 416,726 shoppers out of a total of roughly 4,000,000 unique shoppers with point cards. We impose this criterion because the number of regular shoppers included in the sample stabilizes for shoppers who go to the store on at least a weekly basis, on average ($k \geq 4$).[14]

We make two additional sample restrictions at the store level to obtain our final sample of 511 stores. First, we drop stores from the sample that exit the panel prior to the last month of our sample. Such stores might exit either because the store closes for business or because management chooses to stop providing data to the marketing firm. Second, we restrict to

[13] To our knowledge, the only other paper using these data is Shoji (2020), who investigates consumer stockpiling behavior in advance of a consumption tax hike in 2014.

[14] We obtain qualitatively and quantitatively similar results for the optimal frequency when we deviate from our baseline sample restriction of $k \geq 4$. See Appendix B.2 for the full relationship between the number of unique shopper IDs and trip frequency. Our spending response estimates are robust to using frequent shoppers *within* each goods subcategory, rather than fixing a balanced panel across all goods.

stores that offer loyalty point card programs. Our results are unaffected when we restrict to a balanced panel of stores which record transactions during all non-holiday dates.

To facilitate comparison of our analysis to other papers in the literature and classify products by durability, we aggregate barcodes into 13 subcategories, plus a category called "raw foods" which contains all fresh, non-packaged food items in our sample.[15] Overall, our sample restrictions generate a set of shoppers who make frequent store visits and conduct most of their grocery shopping at stores we can observe in our data. Roughly 28% of the shoppers in our sample have reached the normal retirement age of 65 as of the beginning of the sample, and the average consumer makes 9 trips to a store each month, with an average of 3 days between each visit.

Family Income and Expenditure Survey (FIES) data. We use the public use data files from the Japanese Family Income and Expenditure Survey (FIES) to gauge how much of the consumption profile of the typical consumer we measure with our retail scanner data. The FIES is similar to the Consumer Expenditure Survey (CEX) in the U.S., and features a diary component in which households included in a nationally representative, six-month rotating panel record their daily spending on grocery and non-grocery goods and services with the assistance of an official interviewer who evaluates the responses twice per month. The publicly downloadable files are anonymized and aggregated to the monthly frequency.

A concern with using retail data to measure spending at high frequency is that we may not completely capture the average individual's consumption profile for two main reasons. The first is that our retail data cover only grocery store purchases. This is not an issue to the extent that we are interested in daily instantaneous consumption, or $C(t)$ in our model environment. In Appendix A.4, we extend our optimal frequency results to scenarios where agents exhibit a preference for commitment, perhaps by withholding consumption to make mortgage payments (Vellekoop 2018), or to save up for lumpy expenditures on durable goods like cars or kitchen appliances (Zhang 2023). The second issue is that even within categories of perishable goods we might underestimate consumption, because shoppers select into our sample by (i) going to one of the stores enlisted with the data provider, and (ii) deciding whether to use a loyalty point card to record transactions.

To address the selection problem, in Appendix B.3, we compare typical monthly spending by subcategory and shoppers' frequency of trips to the store vis à vis FIES monthly spending

[15] We make use of the detailed system of classification codes provided by Magee Co. Each barcode in our retail panel is matched to a set of 1-digit to 4-digit categories; for instance a tomato would fall into the "agricultural products," "vegetables" and "fruits and vegetables" codes. We describe the contents of these categories in more detail and offer summary statistics in Appendix B.1.

averages. While our baseline estimation sample of weekly shoppers underestimates monthly spending relative to the FIES by roughly 50%, most of this undercounting is driven by fresh foods like fruits and vegetables, rather than the splurge goods which we show in Section 4 are the most sensitive to payment receipt. Moreover, across both spending data sources the share of perishables in grocery spending is nearly identical at 45%, regardless of store visit frequency. In Appendix F.1 we obtain quantitatively and qualitatively similar results when we re-estimate our main specifications on a roughly 1% subsample of very frequent shoppers for whom we can match average monthly FIES spending.

Japanese Study of Aging and Retirement (JSTAR) data. Since our primary dataset is at the retail level, we can only observe the small number of characteristics that shoppers report when they sign up for a loyalty point card. To gain a more detailed snapshot of pensioner demographics we examine survey responses to the Japanese Study of Aging and Retirement (JSTAR) conducted by the Research Institute of Economy, Trade and Industry (RIETI). The JSTAR is a biannual panel survey of individuals aged 50 or older that is modeled after the Health and Retirement Study in the U.S. We pool responses from the 2007, 2009, 2011, and 2013 waves of the survey, each of which includes data from interviews with approximately 4,000 individuals across ten municipalities in Japan.

We use the JSTAR data to help interpret our empirical results along two dimensions. First, we determine the distribution of ages at which pensioners begin claiming their benefits. Since we identify payments based on a shopper's age, accounting for the potential endogeneity of claimant ages is critical to our research design. We use survey responses related to elderly retirement decisions to shed light on the drivers of early or late claiming of public pension benefits. Ultimately, we replicate the result in Shimizutani & Oshio (2016) that over 90% of pension recipients begin claiming benefits at or prior to reaching 65 years old. Thus, we use age 65 as our cutoff for payment receipt in the event study analysis. Second, we use detailed questions related to wealth and income to assess whether elderly households are liquidity constrained. Responses to the JSTAR corroborate our findings in the scanner data that much of the spending responses around payday cannot be rationalized by common proxies for liquidity constraints.

Public services expenditures on the elderly. We collect annual statistics from the Cabinet Office Historical Data on municipal expenditures towards elderly welfare and aggregate statistics on administrative costs from annual reports published by the national pension system. Such expenditures include local outlays for means-tested benefits targeting the elderly and any administrative costs the local government incurs from distributing national

pension benefits. While the vast majority of public pension payments are processed through the national office and wired via direct deposit to claimants' bank accounts, the national government outsources day-to-day administrative functions to local branch offices.[16]

We merge these data with a list of locations for 312 branch offices tied to the national pension system which are responsible for processing and mailing benefits. All municipalities have a processing center (i.e. a city hall) that issues pension certificates, but some municipalities ($\approx 14\%$) have a separate branch office which both provides data used to process payments and determines pension eligibility. Therefore, municipalities with a branch office are more likely to bear many of the administrative costs of implementing the pension system that would vary with benefit frequency. By comparing elderly expenditures in locations with a branch office to those without an office, we estimate in Section 6.1 how the government's administrative costs for providing pension payments changed following a 1988 reform in which payments switched from a quarterly to bimonthly disbursement schedule.

3.2 Background on the Japanese Old-Age Pension System

The Japanese National Pension System (JPS) is the largest old-age insurance program in the world by fund assets, with annual spending amounting to $474 billion in 2020, or 9.4% of nominal Japanese GDP (OECD 2021). The system consists of two tiers: (i) a National Pension (NP), a flat-rate basic pension with required contributions for residents aged 20 to 59, and (ii) Employee Pension Insurance (EPI), an earnings-related pension with compulsory coverage for those employed full-time by private companies (public employees receive a similar, albeit separate, earnings-based pension). Although the two JPS tiers are distinct in terms of how they determine contribution amounts, both tiers are implemented together as one system, operating according to the same payment schedule.[17]

Full NP eligibility begins at age 65 (the "normal" retirement age) for those with at least 25 years of coverage under the system. Participants have the option of claiming benefits early starting at age 60 in exchange for a permanently lower annual payment or delaying receipt (possible until age 70) in exchange for a permanently higher annual payment. In 2012, the annual full benefit amount for the flat-rate annuity portion of the JPS was 780,100 JPY ($\approx$ $9,000 in 2012) for those with 40 years of contributions.

[16] Such administrative functions include providing consultations, processing applications, providing residency and payment records to the national office, confirming eligibility, investigating fraud, and reconciling benefits.

[17] Coverage under the NP and EPI is voluntary only for a few groups, including those with very low income, non-resident citizens, and elderly immigrants.

Annual EPI benefits are annuitized via the formula

$$EPI = (A - NP) + B + C \tag{3.1}$$

where NP is the flat-rate annuity the claimant receives under the basic NP tier of the system. A is a fixed amount times the number of contribution months, B is a function of average monthly earnings while employed, and C is an additional allowance for dependents. EPI eligibility begins at 65 but a "bridge" pension equal to $A + B + C$ is available for those aged 60-64. The EPI is subject to an earnings test where payments are reduced or suspended if the sum of the bimonthly EPI payment and wages exceeds a threshold of 460,000 in 2012 JPY ($\approx$ \$4,050).[18] Taken together, these features of the system imply that the vast majority of Japanese begin receiving some benefits from JPS once they turn 60. Tabulations from the JSTAR data in Appendix B.3 confirm this. Over 50% of pensioners in the JSTAR sample report claiming some public pension benefits by age 60 (the early retirement age), and over 95% begin claiming by age 65 (the normal retirement age).

Eligible claimants receive pension payments every two months on the 15th of each disbursement month (February, April, June, etc.). If the scheduled delivery date falls on a Saturday, Sunday, or public holiday, payments are instead sent on the first previous non-holiday weekday. This timing rule, combined with annual variation in calendar weekdays, induces variation in the length of periods in between payments that is unrelated to pensioners' spending decisions, conditional on seasonality and intra-week shopping patterns. Our sample period covers 22 scheduled payments, seven of which were rescheduled due to overlap with weekends or holidays. The average interval length among these pay cycles is 59.8 days, with interval lengths ranging from 57 days to 62 days.

4 Expenditure Response to Payments

In this section we describe our main results and high-frequency difference-in-differences strategy for identifying the effect of pension payment receipt on nominal expenditures.

[18] The system generates some income redistribution within age cohorts through B by down-weighting income earned in the 10 years prior to retirement; this feature limits the extent to which pensions replace income from retirement bonuses. More details on benefit formulas can be found at http://www.nenkin.go.jp/international/english/healthinsurance/employee.html.

4.1 Identification Strategy

In our baseline analysis, we adopt the following event study regression to estimate the effects of pension receipt on expenditures:

$$\frac{X_{i,c,t}}{\overline{X}_{i,c}} = \sum_{j=-7}^{+7} \beta_j \cdot Payment_{i,t+j} + \delta_{dow} + \phi_{wom} + \psi_{my} + \xi_h + \eta_i + \epsilon_{i,c,t} \tag{4.1}$$

where $X_{i,c,t}$ is expenditures of shopper i on goods within category c on date t, and $\overline{X}_{i,c}$ is average daily expenditures of the shopper on goods within the category. δ_{dow} are day-of-week fixed effects, ϕ_{wom} are week-of-the-month fixed effects, ψ_{my} are month-by-year fixed effects, and η_i are individual fixed effects. ξ_h is a dummy equal to 1 if date t falls on a public holiday. The dummy $Payment_{i,t+j}$ is an indicator equal to 1 if shopper i is scheduled to receive a payment at time $t+j$. The day-of-week dummies control for intra-week patterns of spending. Week-of-the-month dummies capture lumpy expenditures that take place regularly within some week of the month.[19] The month-by-year dummies account for business cycle conditions and seasonality. We cluster standard errors at the individual shopper level.

Since we do not directly observe pension receipt in our data, we set $Payment_{i,t+j}$ equal to 1 if the shopper is older than the normal eligibility age of 65 for the national pension and $t+j$ is a scheduled pension delivery day.[20] The coefficients β_j measure the fraction by which a shopper's expenditures deviate from the average daily level of spending in a one-week window around the scheduled pension delivery date. Because we do not directly observe pension receipt, our estimates $\widehat{\beta}_j$ are intent-to-treat (ITT) estimates of the effect of pension receipt on expenditures.[21]

The model in (4.1) is a high-frequency difference-in-differences regression that compares the expenditure paths of pension eligibles to non-eligibles. Our identifying assumption is that there is no omitted variable that would differentially impact eligibles and non-eligibles on

[19] These lumpy expenditures may be due to concurrent deadlines for mortgage, utility, and other bill payments, or retail sales campaigns which offer discounts at certain times of the month. We discuss the possibility of the latter in Section 5. In many specifications, we include chain or chain × date fixed effects to difference out pricing strategies typically set at the regional chain level (DellaVigna & Gentzkow 2019).

[20] We also estimate results using earlier and later age cutoffs for treatment, spanning the earliest (60) and the latest age (70) at which beneficiaries can begin claiming benefits. However, the change in the point estimates is negligible as we increase the age threshold. We find this unsurprising given that in the JSTAR data over 90% of retirees begin claiming prior to age 65, and roughly half begin claiming at age 60.

[21] While some shoppers in our sample that are under age 60 might be eligible for spousal, survivor, or disability benefits which arrive on the same payment date, any expenditure response among the under age 60 population to these types of benefits would bias our estimates downward. We disaggregate treatment effects by age cohort in Appendix F.5 and find that payday responses are negligible for under 60-year-olds.

a scheduled payment date, conditional on time-invariant shopper characteristics and fixed effects that capture typical shopping patterns at monthly, weekly, and daily frequencies. One potential issue is that individuals in each of the two groups of shoppers might differ in their propensity to visit a store on a given date, because non-pensioners are more likely to be working during the week than pensioners and thus conduct more of their shopping on weekends. This is a concern due to the fact that JPS reschedules delivery dates that would otherwise fall on Saturdays or Sundays. Hence, the coefficients β_j may capture the effect of pension receipt plus differences across the treatment and control groups in the propensity to visit a store on a payment date, even if no payment were actually delivered.

We address this concern by augmenting the regression in (4.1) in two ways, with results presented in Appendix F.1. First, we interact the individual fixed effects with the full set of day-of-week dummies to account for intra-week shopping patterns that are specific to each shopper. Second, we introduce a new variable, $Period_{i,c,t}$, defined as the number of days at date t since shopper i last purchased within category c. We then create a set of dummies for each quartile of $Period$ and interact these dummies with individual fixed effects.[22] The interpretation of $Period$ depends on the category of expenditures under consideration. For highly aggregated categories such as all raw foods expenditures or total expenditures, this interaction will capture individuals' propensity to shop on certain days of the month. For less aggregated categories such as salad products, this interaction will instead capture the frequency with which shoppers consume goods in that category. Thus we believe the interaction with day-of-week effects is a more appropriate strategy for analyzing expenditure patterns of specific categories of goods.

Differences between treatment and control groups in store visits could also be driven by retailer-specific pricing responses. For instance, if pension recipients are more likely to favor stores in a particular retailer chain, those stores might offer temporary sales promotions on products in elderly shoppers' expenditure baskets to attract pensioners around paydays. To account for retailer strategies, we compare our point estimates obtained with and without including fixed effects for the 20 grocery chains in our sample. In Appendix D, we examine retailers' temporary sale responses using standard filtering methods for identifying sales (e.g. Nakamura & Steinsson 2008). While we uncover evidence consistent with stores offering fewer and less generous discounts on high-quality goods (measured by a modal price) on paydays, when we consider all goods in the expenditure basket, stores are only slightly more

[22]For instance, if a shopper makes a purchase at time t, but their last purchase was in $t-30$, we set $Period_{i,c,t} = 30$; in $t+1$ we then set $Period_{i,c,t+1} = 1$. Since we always restrict to shoppers who make at least one purchase within a category each month, $Period$ falls between 1 and 30 days, and these quartiles correspond roughly to weeks since last purchase.

likely to offer sales on paydays vs. non-paydays. We present results in Appendix F.1 using expenditures deflated by store-level price indices as the outcome variable, which is a closer analog to the notion of consumption in the model environment of Section 2.1.

Finally, we emphasize that because payment dates are rescheduled when a payday falls on a weekend or holiday, the presence of weekends and holidays generates exogenous variation in the day of the month that income arrives. Our specification isolates this exogenous variation even without interacting individual fixed effects with day-of-month fixed effects because all pension claimants receive benefits on the same day.

4.2 Main Results for Expenditures

Figure 3 plots the event study coefficients β_j and 99% confidence intervals obtained from estimating equation (4.1) for total goods expenditures. For our preferred sample of weekly shoppers (Panel A), there is a clear spike in intensive margin spending equal to 9% of average daily expenditures on the scheduled pension date. There are also positive spikes corresponding to 3-4% of average daily expenditures three days prior to and two days after the scheduled pension date. We speculate that these smaller responses within a few days of the delivery date are due to variation in the timing of banks processing transfers from the government to individual claimants' accounts. Overall, the cumulative response of expenditures within a one week window (3 days before until 3 days after the payment date) is 10% of average daily expenditures. Spending responses are more muted for very frequent shoppers who visit the store at least once every other day (Panel B) – at 6% of average daily expenditures $\pm$ 3 days around payday – but otherwise follow a similar trajectory.[23]

Shoppers largely spend more on payday on certain discretionary categories of goods such as alcohol, pre-packaged or prepared meals, and desserts. Table 1 shows heterogeneity in the response of expenditures when we estimate the event study equation in (4.1) using expenditures on different categories of goods to construct the dependent variable. We observe baseline spending responses ranging from 2.8% (processed fish) to 13.7% (alcohol and tobacco) of average daily expenditures on goods within a category. The result that discretionary categories of grocery spending react more strongly to regular income receipt reflects the notion that a large fraction of pension recipients exhibit payday liquidity, or act as if they have a license to splurge on more infrequently purchased goods.

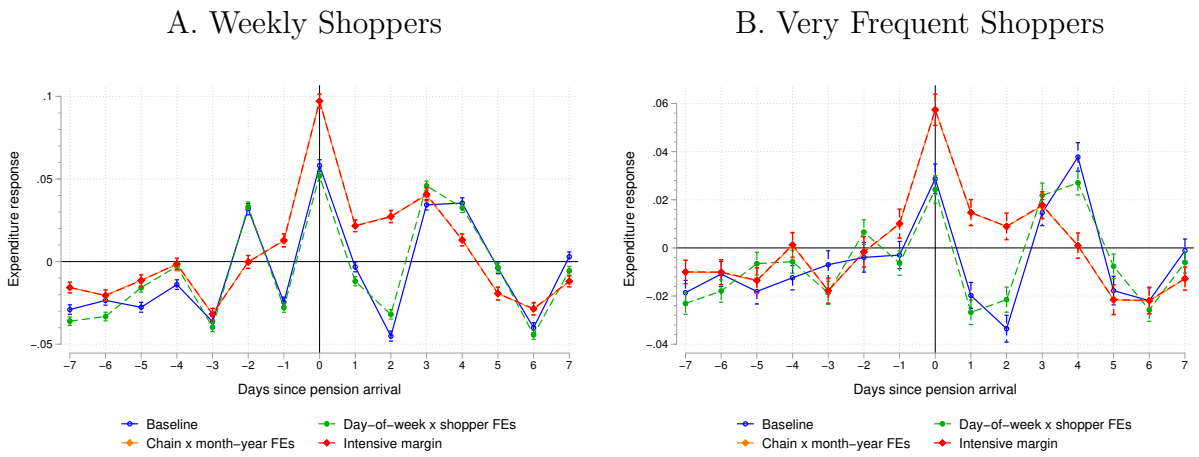
[23] If we recalibrate our structural model according to the spending responses of the very frequent shopper sample, for quasi-hyperbolic consumers we compute an optimal paycycle length of $T^* = 27.33$, compared to $T^* = 27.32$ days for the weekly shoppers we use in our baseline estimation ($\nu = 0.0010$ vs. $\nu = 0.0017$).

TABLE 1. Payday Spending Responses by Margin and Goods Category

Category	Overall	Incl. Chain FEs	Intensive	Extensive
All goods	0.059***	0.099***	0.096***	0.001***
	(0.001)	(0.002)	(0.002)	(0.000)
Raw foods	0.053***	0.093***	0.093***	0.001
	(0.001)	(0.002)	(0.002)	(0.000)
Prepared foods	0.079***	0.212***	0.219***	0.002***
	(0.003)	(0.007)	(0.006)	(0.000)
Sweets/desserts	0.069***	0.166***	0.167***	0.006***
	(0.003)	(0.007)	(0.007)	(0.000)
Alcohol	0.137***	0.275***	0.281***	0.004***
	(0.008)	(0.053)	(0.051)	(0.000)
Fresh produce	0.044***	0.077***	0.076***	0.001*
	(0.002)	(0.003)	(0.003)	(0.000)
Fresh fish	0.060***	0.226***	0.225***	0.001**
	(0.003)	(0.009)	(0.009)	(0.000)
Meat & poultry	0.049***	0.141***	0.132***	0.002***
	(0.002)	(0.005)	(0.004)	(0.000)
Grains	0.024***	0.092***	0.073***	0.001+
	(0.003)	(0.009)	(0.008)	(0.000)
Non-alcoholic beverages	0.048***	0.110***	0.101***	0.005***
	(0.002)	(0.006)	(0.006)	(0.000)
Tobacco	0.137***	0.135	0.140+	0.001
	(0.026)	(0.086)	(0.079)	(0.001)
Processed fruits/vegetables	0.051***	0.180***	0.120**	0.002***
	(0.006)	(0.039)	(0.037)	(0.000)
Preserved fish	0.028***	0.064***	0.060***	0.002***
	(0.003)	(0.011)	(0.010)	(0.000)
Other processed foods	0.056***	0.107***	0.102***	0.003***
	(0.002)	(0.003)	(0.003)	(0.000)

Notes: The table reports results for total spending and intensive and extensive margin versions of regression (4.1) using a panel of weekly shoppers and excluding leads and lags of *Payment*. Each cell in the table is the coefficient on *Payment* from a separate regression within a particular expenditure subcategory. Overall refers to the spending response estimated according to (4.1) and including shopper-day observations of zero expenditures. The second column indicates how our point estimates of the overall spending response changes when we include store chain fixed effects. The dependent variable in the intensive margin regressions is expenditures on a store visit relative to average daily expenditures. The dependent variable in the extensive margin regressions is a dummy for whether the shopper makes a purchase on a given date. In each regression, we winsorize the top 1% of total daily expenditures. Robust standard errors clustered by shopper ID in parentheses. $^{***}p < 0.001, ^{**}p < 0.01, ^{*}p < 0.05, ^{+}p < 0.1$

FIGURE 3. Response of Total Expenditures to Payday



Notes: The figure plots in blue the event study coefficients $\widehat{\beta}_j$ obtained from estimating regression equation (4.1) for all goods expenditures over a fixed panel of either weekly shoppers (Panel A), or very frequent shoppers for whom we can match average monthly grocery spending in the nationally representative FIES (Panel B). Point estimates in green interact the shopper fixed effects η_i with the day-of-week fixed effects δ_{dow} to account for individuals' preferences for shopping on particular days of the week. Point estimates in orange obtained from a version of equation (4.1) which adds store chain $\times$ month-year fixed effects to account for variation in spending patterns due to time-varying retailer-specific pricing decisions. The specification in red is estimated off the intensive margin sample of shopper trips. The y-axis records the percent increase in expenditures relative to panel average daily expenditures. We winsorize the top 1% of total daily expenditures. Bars indicate 99% confidence intervals, with standard errors obtained from clustering by shopper ID.

In even months when payments are delivered, shoppers who receive a payment spend more at the store around the payment date but are not much more likely to make a trip to a store compared to the same date in odd months when no payment is received. In Table 1, to isolate the intensive margin responses to pension receipt, we estimate a version of equation (4.1) defined at the individual-trip level, excluding observations within the panel where we observed zero spending. For the extensive margin, we instead replace the dependent variable in equation (4.1) with an indicator equal to unity if the shopper's expenditures are strictly positive on a given date. For all goods expenditures, the intensive margin estimates indicate that shoppers spent 9.6% more on payday visits relative to average expenditures on other days when they visited a store. The point estimates along the intensive margin are broadly similar across spending categories to specifications in which we include store chain fixed effects. On the extensive margin, the probability that a consumer eligible for pension payments visits a store is roughly 0.1% higher on payday. The results by goods category mirror our event study results; both the intensive and extensive margin responses are more

pronounced for discretionary goods spending.[24]

4.3 Heterogeneity by Pay Cycle Length

The results in the previous subsection on pensioners' expenditure responses to payday reflect average responses across pay cycles with slightly different lengths. Since payment dates are rescheduled when a payday falls on a weekend or holiday, pay cycle lengths in our sample time period range from 57 to 62 days. As noted in our presentation of the payday liquidity model of the consumption path in Section 2.3, there may be more pent-up demand when pay cycles are longer. In other words, the magnitude of the spike in expenditures may be greater when benefit recipients must wait longer in between consecutive payments. This is of direct relevance to the government's problem of setting the optimal payment frequency, because the presence of pent-up demand implies that shorter pay cycles help limit the extent of consumption non-smoothing that is concentrated on payday.

We test for this pent-up demand mechanism by augmenting equation (4.1) with terms that interact the payment dummy with a polynomial function of the pay cycle length:

$$\frac{X_{i,c,t}}{\overline{X}_{i,c}} = \beta_1 \cdot Payday_t \times Length_{t\in p} + \beta_2 \cdot Payday_t \times \left(Length_{t\in p}\right)^2$$
$$+\beta_3 \cdot Payday_t \times \left(Length_{t\in p}\right)^3 + \delta_{dow} + \phi_{wom} + \psi_{my} + \xi_h + \eta_i + \epsilon_{i,c,t} \tag{4.2}$$

where $Length_{t\in p}$ represents the length in days of pay cycle p. It is useful to analyze this regression specification through the lens of the payday liquidity model. When $Payday_t = 0$, it is because i is eligible but t is not a payday. Then in the payday liquidity model payday consumption is determined by the smoothed value: $C_0 = \bar{c}$. We estimate this regression using only the subsample of shoppers over the normal retirement age of 65; otherwise, if we include non-pensioners, the control group will not identify $\bar{c}$ within a pay cycle for recipients. In contrast, when $Payday_t = 1$, equation (4.2) says that payday consumption is some multiple of the smoothed value: $C_0 = \bar{c} \cdot (1 + \beta_1 T + \beta_2 T^2 + \beta_3 T^3) \equiv \bar{c} \cdot (1 + x(T))$.[25]

Table 2 shows from estimating equation (4.2) that shoppers tend to spend more on payday

[24] We plot the event study coefficients by expenditure subcategory in Appendix F.1. We document the same spending patterns regardless of whether we estimate over the sample of weekly or very frequent shoppers

[25] In reported results, we also estimate difference-in-differences versions of equation (4.2) where we include the control group of non-pensioners in the sample, but this barely changes our estimates of $x'(T)$ even if we allow for a spillover effect of $Length_{t\in p}$ on non-claimants. Such spillovers can arise if, for instance, a non-claimant shops on payday and their consumption gets crowded out by the pent-up demand of their claimant spouse who just received a payment.

TABLE 2. Testing the Model of Payday Liquidity with Pent-Up Demand

	(1)	(2)	(3)	(4)	(5)	(6)
$Payday \times Length$	0.0010^{***}	0.0013^{***}	0.0001^{+}	-0.0347^{***}	-1.0550^{***}	-0.1910^{***}
	(0.0000)	(0.0000)	(0.0006)	(0.0010)	(0.0297)	(0.0448)
$Payday \times Length^2$			0.0000	0.0006^{***}	0.0353^{***}	0.0058^{***}
			(0.0000)	(0.0000)	(0.0010)	(0.0015)
$Payday \times Length^3$					-0.0003^{***}	-0.0000^{***}
					(0.0000)	(0.0000)
Time FEs	✓	✓	✓	✓	✓	✓
Intensive margin		✓		✓		✓
$\widehat{x}(T = 60)$	0.060	0.078	0.058	0.078	0.276	0.096
Joint F-test (p-value)	–	–	0.000	0.000	0.000	0.000
N	210,469,638	86,632,913	210,469,638	86,632,913	210,469,638	86,632,913
# shoppers	361,740	361,740	361,740	361,740	361,740	361,740
Adj. R^2	0.025	0.329	0.025	0.329	0.025	0.329

Notes: The table reports results from estimating versions of equation (4.2), which interacts a dummy for pension receipt with the number of days since the last pension payment, for different polynomial orders. The dependent variable is expenditures on raw foods relative to average daily expenditures on raw foods. We use raw foods as the spending category for this exercise because it encompasses perishables which closely approximate consumption. $\widehat{x}(T = 60)$ is the implied fitted value for the payday spike in consumption at $Length = 60$. We report p-values from an F-test for joint significance of coefficients on interaction terms. We include a full set of time fixed effects in all columns, and restrict to the intensive margin of store visits in even columns. Standard errors clustered by shopper ID. $^{***}p < 0.001, ^{**}p < 0.01, ^{*}p < 0.05, ^{+}p < 0.1$

after a longer pay cycle. In the baseline case on the intensive margin where we assume a linear interaction between payment and pay cycle length (column 2), this translates to a 0.13 percentage point increase in the magnitude of the spike per extra day pensioners have to wait for a payment to arrive.

Although the results in columns 4 through 6 of Table 2 provide evidence of non-linearities in pent-up demand following the end of a pay cycle, we note that due to the limited variation in length of the fourteen pay cycles that occur within our sample time period, our polynomial specifications are not globally well-defined. From this exercise, we therefore take the coefficient on the linear interaction $Payday \times Length$ in column 2 to obtain our preferred estimate of the expenditure spike function: $x(T) = 0.0013 \cdot T$. We use this estimate to calibrate the payday liquidity version of our optimal frequency model in Section 6.

4.4 Heterogeneity in Non-smoothing Behavior

The evidence presented so far shows that consumers exhibit a spike in spending on or around the receipt of predictable and regular pension payments. What are the underlying economic mechanisms driving these responses? Our findings are consistent with features of the payday liquidity and quasi-hyperbolic discounting versions of the theoretical framework in Section 2, but other explanations, such as near-rationality or liquidity constraints could play a role.

Distinguishing between these mechanisms is important for the welfare analysis we conduct in this paper. Near-rational consumers induce a welfare loss from excess spending that is proportional to the size of the payment relative to their permanent income (Kueng 2018). Recipients for whom the payment is a large fraction of permanent income will smooth consumption due to the higher potential welfare loss, and those for whom the payment is a small fraction of permanent income will not smooth consumption but incur very small welfare losses. In any case, if pay cycle length is unrelated to the size of the payment relative to permanent income, near-rationality implies the government cannot influence the size of the welfare loss by altering payment frequency.

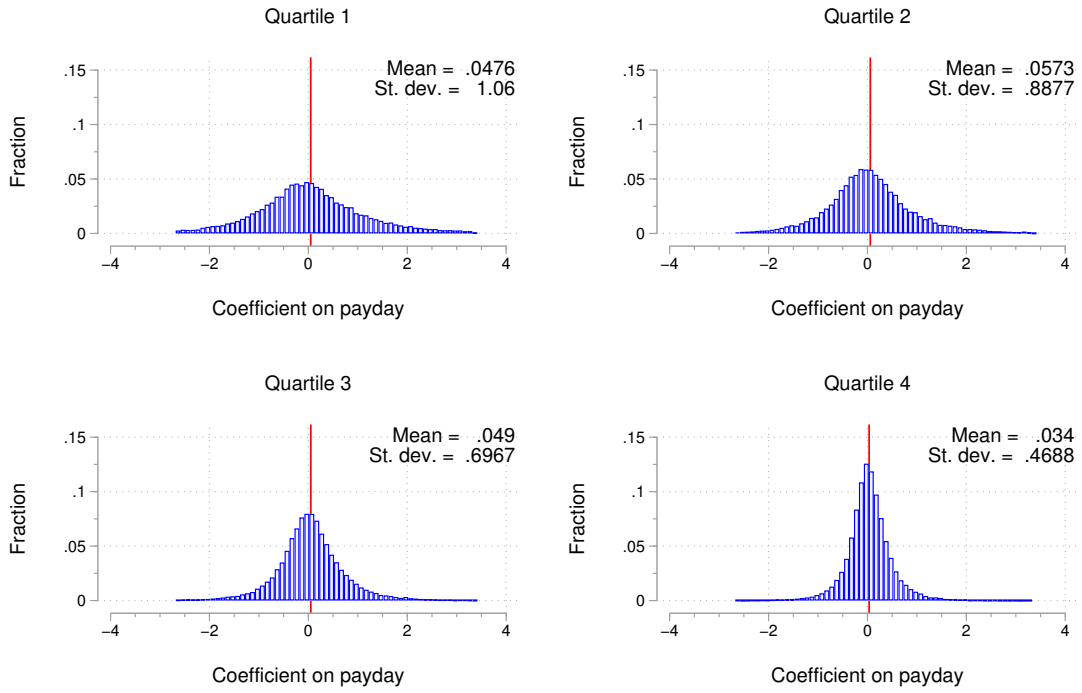
Our framework applies to a regular stream of payments with predetermined amounts. Moreover, in our empirical application, unanticipated variation in consumption needs over a pay cycle is likely to be limited, as universal health insurance coverage in Japan reduces consumption risk from illness and most retirees own their homes. For these reasons, if liquidity constraints exist, they will be highly idiosyncratic in nature. It is therefore difficult to envision how the government's choice of payment frequency could influence welfare by relaxing short-term liquidity constraints.

To further shed light on the motivations for shoppers' spending around payday, we look at heterogeneity in the cross-section of shoppers by average total expenditures over two-month pay cycles in our sample time period. We follow Kueng (2015, 2018) in using total expenditures to proxy for permanent income, as we do not observe other potential proxies for permanent income, such as earned income streams.[26] In the following exercise, we restrict to shoppers aged 65 or over who are eligible for pension payments and who are regular shoppers, defined as those who visit a store at least four times per month, or one week on average. These restrictions leave us with 116,533 shoppers.

We run our main specification in equation (4.1) separately for each of these shoppers using intensive margin raw foods expenditures as the outcome variable and collect the coefficients

[26] We emphasize that in our setting there is limited scope for measurement error from using expenditures as a proxy for permanent income since purchases are not self-reported by households.

FIGURE 4. Payday Responses by Quartile of Pay Cycle Expenditures



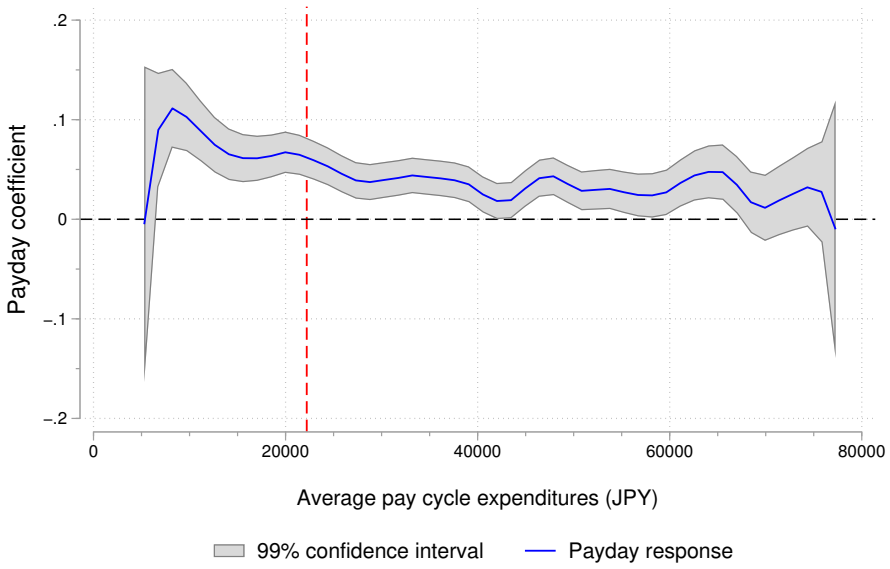
Notes: We sort shoppers by quartile of average total pay cycle expenditures, which we consider a measure of permanent income. We then estimate the time series regression $X^i_{c,t}/\overline{X}_{i,c} = \beta^i \cdot Payday_t + \delta_{dow} + \phi_{wom} + \psi_{my} + \xi_h + \epsilon^i_t$ for each individual shopper ID using raw foods spending, and plot the distribution of the $\widehat{\beta^i}$ by each permanent income quartile. Vertical red lines indicate the within-quartile average payday response. We winsorize the top 1% of daily expenditures.

on the *Payment* dummy. The raw foods category refers to highly perishable goods, so that expenditures within this category are reasonable proxies for instantaneous consumption.[27] Figure 4 plots the distribution of these individual payday responses binned by quartiles of average total expenditures over the pay cycle.[28] The figure shows that spending responses are more variable among low permanent income shoppers; such shoppers visit the store less frequently and have more "lumpy" expenditures than high permanent income shoppers. Yet, the mean response is relatively flat across permanent income bins. Average payday responses are 4.8% in excess of average daily expenditures for the first quartile, 5.7% for the second, 4.9% for the third, and 3.4% for the fourth quartile.

[27]We provide more details on the contents of this category in Appendix B.1.

[28]To limit variance in permanent income due to measurement error, we subset to shoppers who visit a store at least four times each month and appear in the sample for at least six full pay cycles. We also winsorize average daily expenditures at the 5th and 95th percentiles to help guarantee payments are a constant fraction of permanent income across different levels of permanent income.

FIGURE 5. Payday Responses as a Function of Average Pay Cycle Expenditures



Notes: We estimate the time series regression $X_{i,t}/\overline{X}_i = \beta^i \cdot Payday_t + \delta_{dow} + \phi_{wom} + \psi_{my} + \xi_h + \epsilon_t^i$ for each individual shopper ID using raw foods spending. The figure fits a local quadratic function to the relationship between payday responses $\widehat{\beta^i}$ and average expenditures over the two-month pay cycle. We winsorize the top 1% of daily expenditures. 99% confidence intervals represented by the gray shaded area.

Figure 5 shows how payday responses vary continuously with our permanent income measure by plotting a local quadratic of the coefficients on *Payday* against average pay cycle expenditures. There is a non-monotonic, but mostly negative, relationship between spending on payday and permanent income for shoppers with below-median pay cycle expenditures on fresh groceries (22,500 JPY ≈ $225). However, this relationship becomes virtually flat once we look at shoppers with above-median average pay cycle expenditures.[29]

The results here are consistent with a robust finding in the literature that excess spending responses to predictable payments are typically concentrated among low-income individuals (e.g. Zeldes 1989; Broda & Parker 2014). To the extent that the ratio of pension payments to income is relatively constant across the distribution of total expenditures, our findings cast doubt on the near-rationality hypothesis advanced in Kueng (2018), under which we would

[29]In Appendix F.1, we separately re-estimate specification (4.1) for each permanent income decile. Along both the intensive and extensive margin, there is a clear negative relationship between payday responses and permanent income below the median, but a flat relationship above the median.

expect a flat profile of spending responses with respect to total expenditures.[30]

We caution against the interpretation of our results as evidence that liquidity constraints are driving payday expenditures in our setting for several reasons. One reason is that our setting features predictable, lump-sum payments, which most shoppers above the age of 65 have been receiving for many years; the average age among shoppers over 65 in our sample is 71. It is then difficult to argue that recipients lack the ability to build up a buffer to smooth variations over the pay cycle (Fuchs-Schündeln & Hassan 2016). Indeed, in Appendix F.4 we show that after accounting for the staggered timing of retirement eligibility across shoppers, raw foods consumption *increases* rather than decreases after shoppers cross either the early or normal retirement age thresholds. This suggests pensioners have sufficient savings to cover the income loss at retirement, at least within a year after reaching retirement age.

Second, as noted by Parker (2017), behavioral factors such as present-bias and rule of thumb spending could contribute to both low income and excess spending responses to predictable payments. Our findings of heterogeneous responses by permanent income would therefore also be consistent with a higher incidence of behavioral traits such as present-bias and mental accounting among low-income shoppers. Third, in Appendix F.6 to proxy for liquidity constraints we rank consumers on a store quality index based on their preferred retailer and show uniform evidence of mental accounting regardless of whether shoppers visit low or high-quality stores.

5 Do Retailers Capture the Incidence of Payments?

We have demonstrated how pension receipt affects nominal expenditures. However, if retailers engage in price discrimination around scheduled payment dates in anticipation of an increase in demand, the payday coefficients β_j in equation (4.1) will capture changes in prices even when the real expenditure amount stays the same. In this section, we investigate how store-level prices respond to pension payment distribution dates. We construct a price index that isolates the retailer's pricing response from consumer substitution across goods but uncover limited evidence of retailer price discrimination.

[30]The relationship between past earnings and pension amounts is linear for most recipients. There is some concavity in the benefits schedule for recipients who receive large corporate retirement bonuses equal to several years of earnings. Under recent rules, the portion of bonus amounts that counts towards pension benefit calculations is capped at 1.5 million yen per month or 5.73 million yen in any year. Winsorizing average pay cycle expenditures helps ensure that payments are a constant share of income in our sample.

5.1 Store-Level Price Indices

We first consider daily geometric average prices within a store:

$$\Phi_{s,t} = \frac{1}{n_{s,t}} \sum_{k} \log p_{k,s,t} \tag{5.1}$$

where $p_{k,s,t}$ is the price for good k at store s on day t, and $n_{s,t}$ is the number of goods sold at store s at day t. We compute the price of individual good $p_{k,s,t}$ by dividing its total sales by total quantity purchased by shoppers over age 65 at store s at day t.[31] Thus, this price index $\Phi_{s,t}$ represents the average (log) price of goods that consumers over age 65, and who are thus eligible to receive pension payments, buy on each day in each store.

To estimate the effect of pension receipt on the average price, we run the following daily event study regression:

$$\widetilde{\Phi}_{s,t} = \sum_{j=-7}^{+7} \gamma_j \cdot Payday_{t+j} + \delta_{dow} + \phi_{wom} + \psi_{my} + \xi_h + \eta_s + \epsilon_{s,t} \tag{5.2}$$

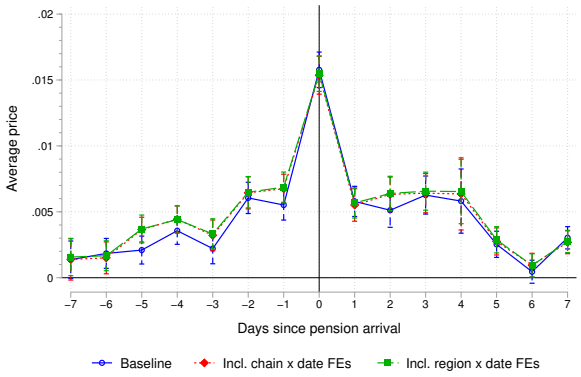
where $\widetilde{\Phi}_{s,t}$ is the daily geometric average price index at store s, normalized by its mean value over our sample period. Hence, $\widetilde{\Phi}_{s,t}$ captures in percentage terms how much store-level prices at date t deviate from their average geometric price index level. As in the expenditure regressions described in Section 4.1, δ_{dow} are day-of-week fixed effects, ϕ_{wom} are week-of-the-month fixed effects, ψ_{my} are month-by-year fixed effects, and η_s are store fixed effects. ξ_h is a dummy equal to 1 if date t falls on a public holiday. $Payday_{t+j}$ is a dummy variable that takes 1 if pension payments are scheduled at time $t+j$. The payday coefficients γ_j are interpreted as how the daily average price of goods purchased within a store deviates from its sample mean if consumers receive pension payments at time $t+j$.

Figure 6 plots the coefficients on the payday dummies γ_j for a two-week interval around pension arrival dates. On payment dates, the average store-level price tends to be 1.6% higher than its sample mean. There are also small but precisely estimated positive effects around pension payment dates. A change in the geometric average price between time t and $t+1$ consists of a change in the number of goods purchased and a change in prices of goods purchased. Panel B shows estimates of equation (5.2) where we replace prices on the left-hand side with the number of goods sold $n_{s,t}$ normalized by its mean value over our sample period. On payday, the number of goods purchased increases by 6.2%. The

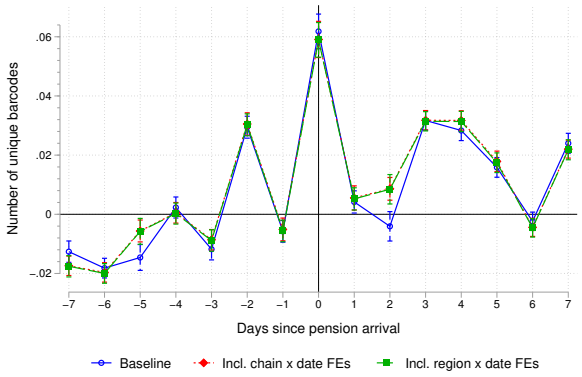
[31] The prices $p_{k,s,t}$ included in our sample are unit prices in the sense that the same product sold in different quantity will be attached to a distinct barcode k.

FIGURE 6. Store-level Price Index and Variety Responses to Payday

A. Average Price Index



B. Number of Unique Goods Purchased



Notes: Panel A of the figure plots the event study coefficients $\widehat{\gamma}_j$ obtained from estimating regression (5.2). The y-axis records the percent increase in store-level prices around payday, using the average price index defined in (5.1). Panel B plots the estimates from (5.2) where we use as the dependent variable the log of the daily number of goods sold (identified by unique barcodes) in a store relative to the average number sold over the sample time period. Point estimates in red obtained from augmenting (5.2) to include store chain $\times$ month-year fixed effects, while estimates in green result from including Census region $\times$ month-year fixed effects. Bars indicate 99% confidence intervals, with standard errors clustered at the store level.

results are impervious to the inclusion of store chain by month-year and Census region by month-year fixed effects, indicating that these patterns are not driven by retailers' seasonal sales promotions or regional shocks such as fluctuations in supply chain resilience.[32]

The positive response of store-level prices to payday could be due to stores raising their prices around pension payment dates, consumers substituting towards higher quality goods when they receive a payment, or some combination of the two responses.[33] If consumers tend to buy higher quality goods around pension paydays, average prices paid by consumers could increase without any change in prices of individual goods set by stores to capture the incidence of payments. In the next subsection, we introduce a simple price index that isolates the change in within-store average prices due to consumer substitution across goods.

[32]Stores in our sample vary in the fraction of regular pensioner shoppers, and so one might expect stores clustered in areas with more pension recipients would be more likely to price discriminate. The fact that our estimates of store-level pricing and sales patterns are insensitive to chain and region-specific trends echoes the findings in DellaVigna & Gentzkow (2019) that chains do not price to market due to managerial inertia and concerns about alienating consumers.

[33]In Appendix F.2, we formally decompose changes in store-level prices into retailer price responses, variety, and substitution effects.

5.2 Price Discrimination or Consumer Substitution?

We consider a counterfactual where there is no retailer pricing response in even months in which pension payments are delivered. In each even month, we replace the prices of individual goods $p_{k,s,t}$ by their last observed daily value in the preceding odd month. We then fix these individual prices over time within the adjacent even month. By fixing prices to be equal to prices in odd months – so that time variation in the price index in payment months only comes from a change in the set of goods purchased by consumers – we can isolate a change in the price index due to consumers' substitution across goods from a change due to a retailer response such as temporary sales.

Figure 7 plots the estimates of the coefficients on *Payday* obtained from equation (5.2) using observed prices minus the same coefficients from using the store-level counterfactual price indices as the dependent variable. The differences are positive for each day within a two-week window around paydays, implying consumers would have paid lower prices if there were no retailer pricing response. However, these differences between the actual and counterfactual pricing responses around payday are economically insignificant. Actual prices are 1.58% higher on payday, while our baseline counterfactual prices are 1.40% higher. The 0.18 p.p. gap between the two estimates captures the retailer's pricing response, which accounts for only 11.39% of the overall spike in payday prices (Panel A).

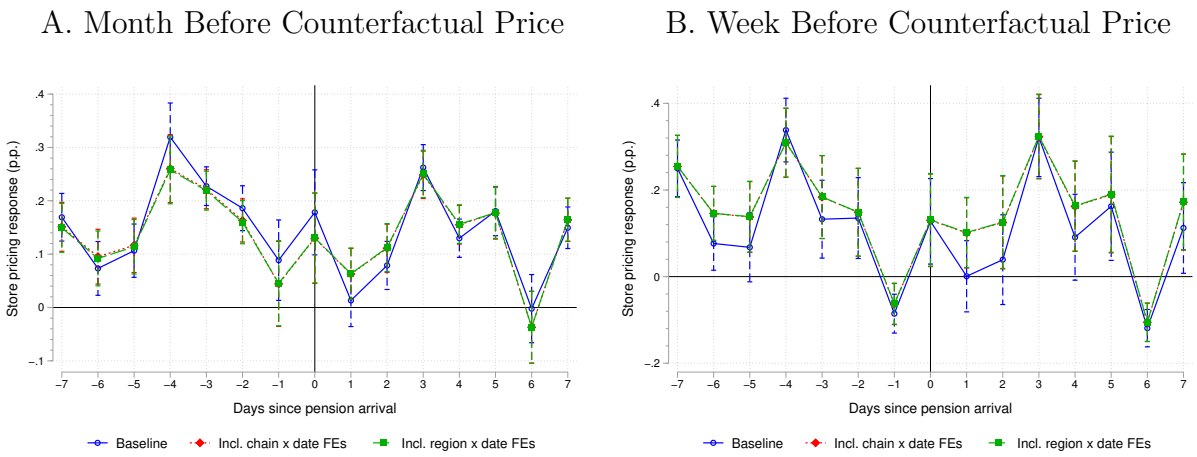
Similarly, for an alternative counterfactual where we use last observed prices exactly one week before payday (Panel B), we compute only a 0.13 p.p. gap, or just 11.40% of the payday spike in the average price index.[34] We provide a more complete analysis of store pricing responses in Appendix F.3 and conclude that for most goods subcategories there is no statistical difference between the average price index and the counterfactual on payday. Thus, while retailers capture a small portion of the incidence of pension payments, the observed increase in average prices around payday is almost entirely due to consumer substitution towards goods of higher value.[35]

In Appendix D, we explore whether retailers engaging in more aggressive sales promotions

[34]The week before counterfactual has the advantage of holding fixed the day of the week and the month-year in comparing the payday price index to the last price. We use the preceding odd month as our baseline counterfactual because the week before one might reflect forward-looking behavior by retailers trying to capture the incidence of payments, which downward biases our estimates of the pricing response when we do not include chain fixed effects.

[35]In unreported results, we show high-quality stores, measured by average prices charged outside payday weeks for regularly-transacted goods, are more likely to price discriminate. Stores in the highest quartile of the quality distribution account for the entire 0.13 p.p. gap between actual and last observed prices on payday.

FIGURE 7. Store Pricing Responses around Payday



Notes: We construct store-level price indices as in equation (5.1) and then estimate regression (5.2) to obtain estimates $\widehat{\gamma}_{1,j}$. We then construct a counterfactual last price index, using either the last observed price for each barcode in the odd month preceding the payday (Panel A) or the last observed price in the week prior to payday (Payday B), and rerun (5.2) on the exact same sample using this price index as the dependent variable to obtain $\widehat{\gamma}_{2,j}$. The figure plots the point estimates and 99% confidence interval bars for the differences $\widehat{\gamma}_{1,j} - \widehat{\gamma}_{2,j}$, with standard errors in each regression clustered at the store level. Point estimates in red obtained from augmenting each version of (5.2) to include store chain × month-year fixed effects, while estimates in green result from including Census region × month-year fixed effects. The differences capture the percentage point increase in prices on payday that cannot be explained by changes in the composition of goods purchased.

on paydays explains these substitution and variety effects. Consistent with an established industrial organization literature on the seasonality of markdowns (e.g. MacDonald 2000; Chevalier, Kashyap, & Rossi 2003), we show that stores are slightly more likely to offer temporary sales on paydays, driven by higher discounts offered on *below-median* price barcodes within a four-digit goods category. Together with the results on consumer substitution towards splurge goods on payday, our evidence supports the explanation in Chevalier & Kashyap (2019) that retailers can attract more price-sensitive "bargain hunters" by offering discounts during peak demand periods. Overall, we uncover limited evidence of price discrimination in response to customers receiving payments, which is consistent with more recent evidence from Goldin, Homonoff, & Meckel (2022) on retailer responses to SNAP in the U.S.[36] To the extent some price discrimination does occur, in Appendix A.5, we incorporate temporary sales into our optimal frequency framework and discuss how pricing

[36]Our failure to uncover substantive price discrimination is unlikely to be due to imperfect targeting of goods on store's side – that is, raising prices on goods with large expenditure shares in pensioners' baskets. Stores appear in our dataset if they contract with a marketing firm which offers them the necessary information to engage in this type of behavior.

responses to payday alter the government's choice of optimal pay cycle length.

6 Model Calibration

In this section we describe how we calibrate our optimal payment frequency model using our estimates of the expenditure response to payments in the preceding sections and statistics on costs and benefit expenditures from JPS.

6.1 Estimating Administrative Costs

Our identification of the slope of the administrative cost function in our model $\mu'(T)$ relies on a 1988 reform to the pension system which reduced the length of the pay cycle from three months to two months.[37] In February 1988, poor elderly recipients of the Old Age Welfare Pension benefit and existing JPS claimants who only qualified for the national pension but not the employee-based pension benefit started receiving benefits in each even month. Hence, the bulk of payments from JPS were switched from a quarterly to bimonthly schedule starting in FY 1987. Since the 1988 reform did not result in any new changes to the formula determining overall benefit amounts, this shift in policy represents a pure shock to T that did not shift the intercept of the administrative cost function $\mu(\cdot)$.

While the policies governing the national pension system are set at a national level, many of the day-to-day functions of implementing the system are delegated to local town governments and local branch offices of the system which tend to be located in municipalities which lie in the center of a commuting zone. In particular, while claimants can apply to start receiving benefits at their local town hall office, the closest branch office is responsible for processing applications, reconciling benefits, confirming eligibility, and investigating fraud. This fiscal federalism built into the system implies that local governments with a branch office within their jurisdiction were more exposed to the shock to administrative costs, and thus would have seen expenditures allocated to the elderly welfare rise more relative to non-branch office governments following the reform.

Given this natural experiment, we run a standard difference-in-differences regression:

$$\log \mu_{j,t} = \beta \cdot Branch_j \times Post_t + \gamma_j + \delta_t + \epsilon_{j,t} \tag{6.1}$$

[37] In Appendix E, we discuss the details related to the data used in this section, as well as our procedures for identifying pension branch office locations.

where $Branch_j$ is a dummy equal to 1 if municipality j contains a branch office of the Japanese National Pension System, $Post_t$ is a dummy equal to 1 if year t is FY 1987 or later, γ_j are municipality fixed effects, and δ_t are (fiscal) year fixed effects. The dependent variable is the log of costs per person over 65 associated with providing elderly welfare benefits, which includes costs in support of the pension system. We consider this measure to be good a proxy for administrative costs per claimant, since over 90% of individuals over age 65 receive benefits from JPS but only a small share of this population receive other forms of welfare benefits. Under the assumption of parallel trends in our cost per claimant measure, β captures the mid-run increase in costs due to increasing the frequency of benefits from every three months to every two months, holding fixed the formula for benefit amounts.

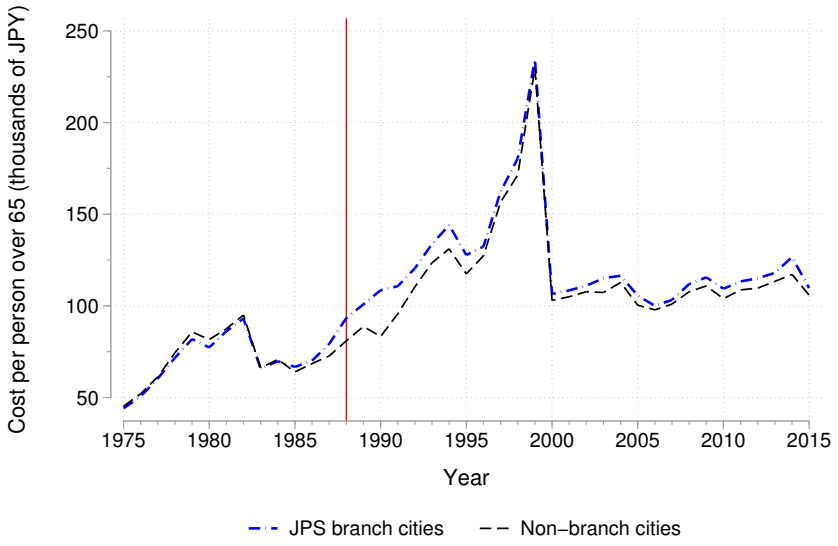
Figure 8 plots the average of this cost measure across all cities with a JPS branch office along with the average across all non-branch cities. The level and slope of costs across the two groups of cities is very similar up until 1987 when the government announced the reform.[38] The cost series continue to diverge until the government froze increases in contribution rates for several years starting in December 1998. The government then proposed an overhaul of the system in 1999 (passed in 2000) which aimed to reduce costs and aggregate spending and benefits.[39] To isolate variation in municipal administrative costs due to the 1988 reform which changed only the payment frequency, we restrict our sample to the years 1980-1996, which creates a symmetric window around the reform.

Table 3 presents results from estimating equation (6.1). Administrative costs increase by approximately 4.3% in branch office cities in the post-reform years, or 7.3% conditional on non-parametric trends by municipal population and income bins. The effect we uncover in columns (1) and (7) disappears once we exclude the 23 governments in metropolitan Tokyo and the five largest cities outside Tokyo from our sample. We speculate that JPS branch offices in central Tokyo were impacted more than other large cities with branch offices due to the comprehensive public transport system reducing the commute time to these offices. In short, we do not uncover evidence that the switch from a quarterly to bimonthly payment

[38]Since the municipal budget data are for a fiscal year, which runs from April 1 to March 31 in Japan, the divergence in the cost series beginning in 1987 is due to overlap of the 1987 fiscal year with the start of the new payment schedule in February 1988, rather than an anticipation effect.

[39]Major tenets of the 2000 reform included switching from wage-indexing to CPI-indexing of benefits, introducing an earnings test for claimants aged 65 to 69, reducing earnings-related benefits by 5%, and gradually increasing the normal retirement age from 60 to 65 (Takayama 2001).

FIGURE 8. Municipal Government Expenditures on the Elderly



Notes: The figure plots time series of municipal spending (in thousands of real 2012 JPY) on administering the pension system and elderly welfare benefits divided by the number of persons over age 65 residing in the municipality. The blue dashed line refers to the average of this ratio over all municipalities with a JPS branch office (N = 239), while the black dashed line averages over municipalities without a JPS branch office (N = 424). The vertical red line indicates the year (1988) when JPS switched from a quarterly to bimonthly payment schedule. We use branch office locations as of 1980 and exclude from the sample municipalities with gaps in the cost data. See Appendix E for details.

schedule had a large effect on government costs of administering the public pension system.[40]

In our optimal frequency model, the government's administrative cost $\mu(T)$ is a function of the number of days T between payments. Taking our baseline estimate of a 4.3% increase in costs associated with a decrease in the pay cycle from 90 to 60 days implies a 0.14% increase in costs for each day the government shortens the pay cycle, assuming $\mu(T)$ is linear on the interval $T \in [60, 90]$. To match our aggregate statistics on the 2012 pension system and calibrate our model, we translate this estimate to real 2012 expenditures per claimant. Scaling reported total administrative costs for FY 2011 by our 4.3% estimate implies an increase in 8,111 JPY per claimant. The implied annual cost for each day the pay cycle length is shortened amounts to 270 JPY per claimant. For our upper bound estimate of a 7.3% increase in administrative costs, the implied cost increase is 13,770 JPY per claimant,

[40]We acknowledge one drawback to this approach is external validity of the results to our sample time period of 2011-2014 for the retail scanner data. Prior to 2008, JPS did not offer internet services, so participants had to visit a local JPS office or a branch office to check their contributions and balances. The slope of the administrative cost function may not be a time-invariant policy parameter if the introduction of internet services reduced the sensitivity of administrative costs to payment frequency.

TABLE 3. Effect of Pension Frequency Reform on Municipal Admin Costs

	(1)	(2)	(3)	(4)	(5)	(6)
$Branch \times Post$	0.043**	−0.012	0.034	−0.003	0.073***	0.021
	(0.017)	(0.014)	(0.023)	(0.016)	(0.026)	(0.017)
City & year FEs	✓	✓	✓	✓	✓	✓
Incl. Tokyo	✓		✓		✓	
Incl. major cities	✓		✓		✓	
1985 population bin × year FEs			✓	✓	✓	✓
1985 per capita income bin × year FEs					✓	✓
N	11,111	10,635	11,111	10,635	11,111	10,635
# Municipalities	663	635	663	635	663	635
Adj. R^2	0.517	0.554	0.856	0.863	0.863	0.866

Notes: The dependent variable in each regression is log expenses on elderly welfare per resident at or above age 65. $Branch_j = 1$ if municipality j contains a Japan Pension System branch office. $Post_t = 1$ for years 1988–1996. All regressions include observations for years 1980 – 1996 and a full set of year fixed effects. Robust standard errors clustered at the municipality level in parentheses. Tokyo consists of the 23 central wards for which separate expenditure time series are available. Major cities consist of the historically five most populous cities outside of Tokyo: Yokohama, Nagoya, Kyoto, Osaka, and Kobe. 1985 population bin refers to quintiles of 1985 Census population. 1985 per capita income bin refers to quintiles of per taxpayer taxable income in 1985. $^{***}p < 0.01, ^{**}p < 0.05, ^{*}p < 0.1$

or 459 JPY per claimant for each day the pay cycle length is shortened.[41]

6.2 Calibration Procedures

We now compute the optimal payment frequency by combining our empirical results on the behavior of consumption expenditures within the pay cycle with our results on the shape of the government's administrative cost function for providing pension benefits. This payment frequency is the solution to the government's optimization problem in equation (2.6) and characterized by the first-order condition in (2.7). We perform computations for the two main versions of our model featuring quasi-hyperbolic discounters and payday liquid consumers. In the version with payday liquid consumers, we assume that the observed spike in consumption on payday is an increasing function of the pay cycle length T, as modeled in equation (2.16).

[41] A simple non-parametric differences in means between the branch office and non-branch office cities pre- vs. post-reform yields an estimate of an increase of 14,464 JPY (in 2012 yen) per claimant due to moving to a bimonthly payment schedule, which implies an increase in 482 yen per claimant per pay cycle day. The absolute difference from our measure based on scaling up FY 2011 administrative costs is primarily due to the larger population of claimants in 2011-2012 relative to the 1980s and 1990s.

This assumption is consistent with our empirical results in Section 4.3, where we showed that the spike in expenditures on payday increases by 0.13 percentage points for each additional day in the pay cycle.

Our modeling framework consists of two sets of parameters: preference parameters captured by ρ in the isoelastic utility case and the deviation function $f(t)$; and parameters associated with the government budget captured by p, $\overline{B}$. We analyze how the behavior of the welfare loss and optimal frequency varies with ρ and any latent parameters of $f(t)$ for each version of our model outlined in Section 2. To calibrate the government's budget constraint, we rely on information on aggregate flows and participants in FY 2011 from the Japanese Pension System Annual Report 2012.[42]

For the administrative cost function, we assume a convex cost function of the form $\mu(T) = \kappa_\ell / T^\ell$ and for each power ℓ calibrate κ_ℓ so that $\mu(60)$ equals the administrative service costs reported for FY 2011 in the annual JPS Business Report.[43] In FY 2011 there were 67.37 million contributors to the system and 38.67 million pension recipients, implying that pensioners made up $p = 0.377$ of all participants. Over the entire fiscal year total pension payments amounted to 48,867.5 billion yen, or roughly 10% of nominal 2012 GDP. This amount implies an average daily payment per claimant of $\overline{B} = 3{,}462$ yen ($\approx$ \$32.50), or an average payment of 207,732 ($\approx$ \$1,950) over the pay cycle. JPS reports administrative service costs of 300.722 billion yen in FY 2011.

Our results in Section 6.1 from exploiting differential exposure of local governments to the 1988 pension schedule reform suggest that $\mu(T) = \kappa_\ell / T^\ell$ with $\ell = 0.27$, calibrated to the aggregate system flows in 2012, is a good approximation for the administrative cost function.[44] While we view this version of the government's cost function as empirically plausible, we also provide results for more or less convex cost functions by altering ℓ to obtain conservative upper bound estimates for the optimal frequency.

Lastly, we choose preference parameters that match the results from estimating expenditure regressions of the form in equation (4.1). We found that expenditures increase by roughly

[42]The official pension system report is available at http://www.nenkin.go.jp/files/13nenkinD_synthesis.pdf.

[43]A copy of the FY 2011 JPS report can be found at http://www.nenkin.go.jp/info/disclosure/jigyo.files/23-2.pdf.

[44]More concretely, we set $\mu(90) - \mu(60)$ equal to the cost per claimant increase implied by our regression estimates in Table 5. The formula used to obtain this value is $\ell^* = \frac{\log(1-\widehat{x}\cdot n/c)}{\log(2/3)}$, where $\widehat{x}$ is the cost per claimant increase implied by our regression estimates, n is the number of pension recipients, and c is total administrative costs reported by the government. Using our preferred estimate of a 4.3% cost increase from the 1988 reform, we obtain $\ell = 0.27$. For our upper bound estimate of a 7.3% cost increase from the 1988 reform, we obtain $\ell = 0.28$.

10% for all goods or all raw foods expenditures upon payment receipt. In the payday liquidity model, this corresponds to a value of $x = 0.1$, which is one-third the estimate of $x = 0.3$ obtained by Olafsson & Pagel (2018) in their study of total expenditures on consumption and non-consumption goods. The estimates are consistent to the extent that grocery store items are less likely than other non-grocery categories of expenditures to be classified by consumers as splurge goods. We use our results on heterogeneity by pay cycle length in Section 4.3 to set $x'(T) = 0.0013$, reflecting that the payday spike in expenditures tends to be 0.13 percentage points higher for each additional day in pay cycle.

Our estimate of a 10% spike in perishable grocery expenditures on payday also allows us to calibrate the degree of present-bias behavior in the version of the model with quasi-hyperbolic discounters. For quasi-hyperbolic discount factors β close to one, the decline in consumption over the pay cycle is approximately linear. Hence, to calibrate the overall daily discount rate in the model, we can set a linear rate of negative consumption growth between time $t = 0$ (payday) and time $t = T - 1$. The fact that raw foods expenditures increase by 10% between the day before the new pay cycle starts and the day a new pay cycle begins implies that this daily rate of negative consumption growth is $\nu = 0.10/60 \approx 0.0017$, or 0.17% of payday consumption. This estimate is half of the $\nu = 0.004$ estimate found by Shapiro (2005) in his analysis of food stamp recipients' consumption between monthly pay cycles, which suggests that present-bias plays a less prominent role in our sample of pension recipients.

6.3 Optimal Payment Frequency Results

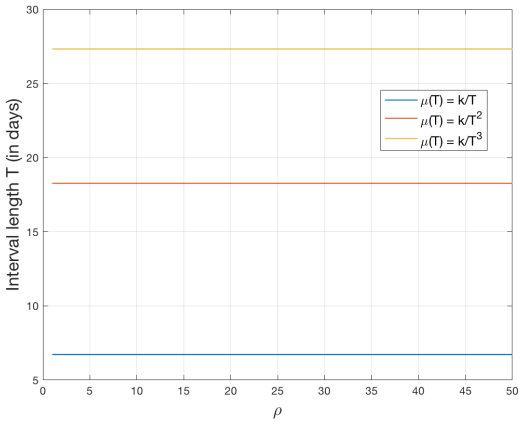
Figure 9 presents the optimal payment frequency for each version of our model, with different values of the inverse IES $\rho \in [1, 50]$ and for three different versions of the government's convex administrative cost functions: $\mu(T) = \kappa_\ell / T^\ell$ for $\ell = 1, 2, 3$. The figure showcases two key results. First, in each model version, the optimal frequency is almost completely invariant to the assumed value for the inverse IES, and the optimal frequency is (weakly) decreasing in ρ. Intuitively, ρ determines the size of the welfare loss from consumption non-smoothing but not the increase in marginal costs to the government of decreasing T. However, the marginal welfare loss $-\lambda'(T)$ is largely invariant to the inverse IES. The optimal pay cycle length T^* is decreasing in ρ because the marginal benefit to the government of decreasing T is steeper for higher ρ. This is because a higher ρ means that consumption is less substitutable between periods, so individuals would be willing to pay more to get closer to the smooth consumption path in the *ex ante* sense in which we have defined the welfare loss.

Second, for each of the administrative cost functions across both model versions, the implied optimal frequency is less than one month. For $\ell = 1$, the optimal frequency is

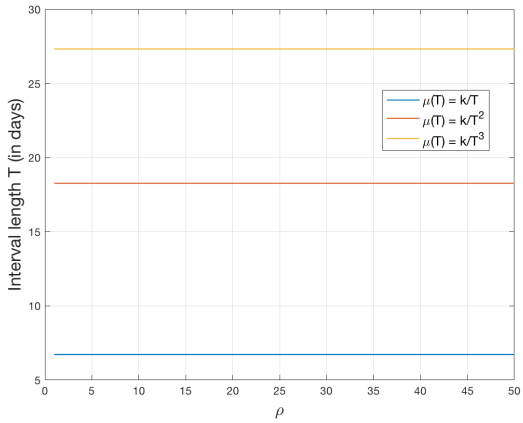
FIGURE 9. Results from Calibrated Optimal Frequency Model

A. Optimal frequency by Inverse IES

Quasi-hyperbolic Discounting

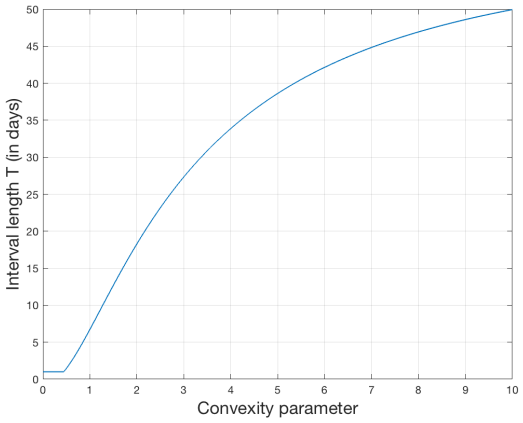


Payday Liquidity

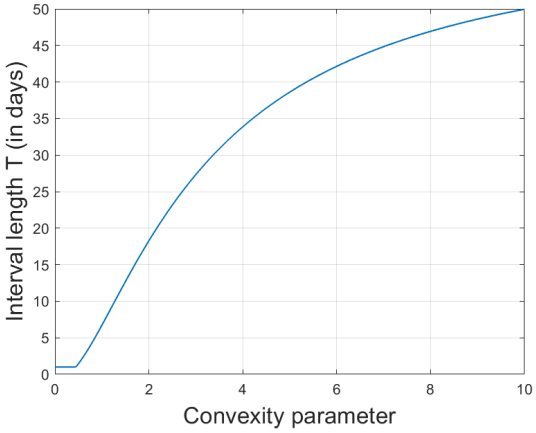


B. Optimal frequency by Convexity of Administrative Costs

Quasi-hyperbolic Discounting



Payday Liquidity



Notes: The figure plots the optimal frequency implied by quasi-hyperbolic discounters and by payday liquid consumers as a function of either the consumers' inverse intertemporal elasticity of substitution (Panel A) or the convexity of administrative costs (Panel B). We solve for T in the government's FOC in equation (2.7) for each version of the model using standard non-linear optimization packages. See main text for details on calibration and Appendix A for derivations.

$T^* = 6.72$ days, whereas for $\ell = 2$ we obtain $T^* = 18.27$ days, and for $\ell = 3$ we obtain cycles of length $T^* = 27.32$ days. In Panel B, we show how T^* varies continuously as a function of the convexity of administrative costs for a fixed value of the inverse IES ($\rho = 10$).[45] For very non-convex cost functions ($\ell \leq 0.45$) costs become negligible and the government finds it optimal to disburse daily payments. Even for extremely convex administrative costs the optimal interval length remains below two months. Thus, the government can improve welfare by distributing pension payments more frequently than the current bimonthly schedule.

6.4 Implications for Pension Eligibility Reforms

Pension reforms proposing to raise the eligibility age have gained considerable attention in developed economies due to increasing life expectancy and resulting financial strain on social security systems.[46] How sensitive is our estimate of the optimal frequency to changes in the eligibility age? This sensitivity analysis is essential given the increasingly common trend towards population aging in developed countries.

In our model, an increase in the pension eligibility age would lead to a decrease in the parameter p, which is the fraction of pensioners in the total population. The baseline estimate under the eligibility age of 65 is $p = 0.377$. When the normal retirement age increases to 70, as Japan declared in April 2021, the new value would be $p^{new} = 0.307$ if the early claiming rate of 10.8% persists for those in the 65-70 age group.[47] Rewriting the FOC of the government's problem in equation (2.7), under the new eligibility age, we have:

$$\frac{\lambda'\left(T^{*new}\right)}{\gamma} = \frac{1}{p^{new}} \cdot \mu'\left(T^{*new}\right) + \overline{B} \tag{6.2}$$

We previously made the following parametric assumption for the administrative cost

[45] Again, the differences across the two versions of the optimal frequency model are only present after the sixth decimal point. Intuitively, this is because both models produce marginal welfare losses that are very flat with respect to payment frequency.

[46] The U.K. has implemented a phased increase in the state pension age, rising from 65 to 66 since 2020, with plans to further raise it to 67 between 2026 and 2028. Similarly, Germany has been gradually raising the general retirement age from 65 to 67. France has passed reforms to incrementally raise the state pension age from 62 to 64 between 2023 and 2030, which led to widespread protests and strikes. In Japan, changes have been implemented to gradually increase the eligibility age from 60 to 65 for the employee pension, with phased-in stages ranging from 2013 to 2025 for men and from 2018 to 2030 for women.

[47] We obtain this early claiming rate by age cohort from the Japanese Pension System Survey for FY 2011, the first year our retail panel covers. According to the Pension System Survey, pensioners who are less than 70 years old account for approximately 40% of the total number of people claiming public pension benefits. This survey is conducted by conducted by Ministry of Health, Labour and Welfare. Data for FY 2011 can be found at this website.

function: $\mu(T) = \kappa_\ell / T^\ell$. Under the new eligibility scheme, we have

$$\frac{1}{p^{new}} \cdot \mu'(T^*) = \frac{\kappa_\ell}{p^{new}} \cdot \Big(-\ell / T^{*\ell-1} \Big) < \frac{1}{p} \cdot \mu'(T^*) \tag{6.3}$$

Thus, the government's optimality condition becomes easier to satisfy, which pushes the optimal T^* up (λ is declining in T). Intuitively, as the fraction of pensioners declines, the marginal cost of increasing the payment frequency dominates the marginal welfare gain.

Re-estimating our model with p^{new} at the age 70 eligibility cutoff, we obtain slightly longer pay cycle lengths than those we estimated in Section 6.3 for the 2011 regime: $T^* = 7.55$ for $\ell = 1$, $T^* = 19.63$ for $\ell = 2$ and $T^* = 28.81$ for $\ell = 3$. This exercise holds fixed the average payment per claimant $\overline{B}$ and the slope of the cost function $\mu'(T)$. It also assumes that the average marginal welfare loss from consumption non-smoothing does not change with the distribution of pensioner ages, or, in other words that relatively young and old pensioners exhibit similar behavioral responses, on average, to payment receipt. We provide empirical support for the latter assumption in Appendix F.5.

Our optimal frequency framework also highlights how lengthening the time between payments could be an alternative to raising the eligibility age to balance social security budgets, to the extent that administrative costs decline with pay cycle interval length T. For instance, consider the proposed April 2021 reform to the Japanese Pension System which is set to raise the normal retirement age (NRA) from 65 to 70. Assuming a distribution of claiming ages reflected in the recent Japanese Pension Survey from 2021, and ignoring potential increases in revenues from income taxes if people stay in the workforce for longer, we estimate this reform would result in annual savings of 36.12 billion JPY. These savings would be attributed to the application of higher penalty rates to individuals in the 65-70 age group. We offer a full description of the steps through which we arrive at this number in Appendix A.6. Using our difference-in-differences estimates from Section 6.1 to pin down the administrative cost curvature parameter, $\ell = 0.27$, an equivalent cost-saving reform would be to move from a monthly system with $T = 30$ (our upper bound on the optimal frequency) to $T = 45$. In other words, increasing the NRA from 65 to 70 would save the pension system the same amount as moving from issuing payments once every month to once every six weeks.

7 Conclusion

We ask how governments should choose the frequency of public benefits. In our simple model, the government chooses the length of the pay cycle subject to a tradeoff between minimizing

the welfare loss from consumption non-smoothing and incurring increased administrative costs. We apply our model to the bimonthly Japanese pension system and find, under a battery of alternative assumptions about consumer preferences and administrative costs, that the government could improve welfare by moving to a monthly disbursement schedule. Yet, for governments facing convex administrative cost functions and rapidly aging populations, our results imply lowering pension payment frequency may be a budget-preserving, but more politically palatable, alternative to increasing the normal retirement age.

Our framework applies more generally to other public benefits programs. We demonstrate under several empirically plausible models of individual consumption paths that welfare losses from non-smoothing are likely to be small. However, the main lesson here is that for public benefits programs where administrative costs vary minimally with respect to the frequency of disbursement, the government can achieve a welfare improvement by splitting up entitlements into smaller, more frequent payments. Even under extreme assumptions about consumer preferences or administrative costs, the optimal interval length rarely exceeds one month, which supports the worldwide prevalence of monthly payment schedules.

Further, our application to high-frequency retail data sheds light on consumer and retailer responses to the timing of income. Shoppers sharply increase expenditures within two days of scheduled payment dates, but most of this response is concentrated in purchases of discretionary goods such as alcohol, desserts, and prepared foods. These findings accord with models of intertemporal decision-making that feature rule-of-thumb spending or mental accounting. While within-store average prices spike on paydays, this is almost entirely due to consumers substituting towards a basket of more numerous, higher quality goods, rather than retailers raising prices or suspending temporary sales on specific items. Our work thus points to estimating consumers' substitution and variety responses to the timing of income payments as a path for future research on transfer programs, including ongoing debates about inflation generated by stimulus payments remitted during the COVID-19 crisis.

References

Agarwal, S., J. Pan, & W. Qian (2015): "The Composition Effect of Consumption around Retirement: Evidence from Singapore," *American Economic Review: Papers & Proceedings*, 105(5): 426431.

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

Baker, S. (2018): "Debt and the Response to Household Income Shocks: Validation and Application of Linked Financial Account Data," *Journal of Political Economy*, 126(4): 1504-1557.

Baker, S., B. Baugh, & L. Kueng (2021): "Income Fluctuations and Firm Choice," *Journal of Financial and Quantitative Analysis*, 56(6): 2208-2236.

Battistin, E., A. Brugiavini, E. Rettore, & G. Weber (2009): "The Retirement Consumption Puzzle: Evidence from a Regression Discontinuity Approach," *American Economic Review*, 99(5): 2209-2226.

Baugh, B., I. Ben-David, H. Park, & J.A. Parker (2021): "Asymmetric Consumption Smoothing," *American Economic Review*, 111(1): 192-230.

Baugh, B. & F. Correia (2022): "Does Paycheck Frequency Matter? Evidence from Micro Data," *Journal of Financial Economics*, 143(3): 1026-1042.

Baugh, B. & J. Wang (2021): "When Is it Hard to Make Ends Meet?, *mimeo*, University of Illinois at Urbana-Champaign.

Bernheim, B.D., J. Skinner, & S. Weinberg (2001): "What Accounts for the Variation in Retirement Wealth Among U.S. Households," *American Economic Review*, 91(4): 832-857.

Best, M., J. Cloyne, E. Ilzetzki, & H. Kleven (2020): "Estimating the Elasticity of Intertemporal Substitution Using Mortgage Notches," *Review of Economic Studies*, 87(2): 656-690.

Bils, M. & P.J. Klenow (2001): "Quantifying Quality Growth," *American Economic Review*, 91(4): 1006-1030.

Broda, C. & J.A. Parker (2014): "The Economic Stimulus Payments of 2008 and the Aggregate Demand for Consumption," *Journal of Monetary Economics*, 68: S20-S36.

Browning, M. & M.D. Collado (2001): "The Response of Expenditures to Anticipated Income Changes: Panel Data Estimates," *American Economic Review*, 91(3): 681-692.

Browning, M. & T.F. Crossley (2001): "The Life-Cycle Model of Consumption and Saving," *Journal of Economic Perspectives*, 15(3): 3-22.

Chetty, R. & A. Szeidl (2007): "Consumption Commitments and Risk Preferences," *Quarterly Journal of Economics*, 122(2): 831877.

Chevalier, J.A., A.K. Kashyap, & P.E. Rossi (2003): "Why Don't Prices Rise During Periods of Peak Demand? Evidence from Scanner Data," *American Economic Review*, 93(1): 15-37.

Chevalier, J.A., A.K. Kashyap (2019): "Best Prices: Price Discrimination and Consumer Substitution," *American Economic Journal: Economic Policy*, 11(1): 126-159.

Cochrane, J.H. (1989): "The Sensitivity of Tests of the Intertemporal Allocation of Consumption to Near-Rational Alternatives," *American Economic Review*, 79(3): 319-337.

Coibion, O., Y. Gorodnichenko, & D. Koustas (2021): "Consumption Inequality and the Frequency of Purchases," *American Economic Journal: Macroeconomics*, 13(4): 449-482.

DellaVigna, S. & M. Gentzkow (2019): "Uniform Pricing in U.S. Retail Chains," *Quarterly Journal of Economics*, 134(4): 2011-2084.

Dobkin, C. & S.L. Puller (2007): "The Effects of Government Transfers on Monthly Cycles in Drug Abuse, Hospitalization and Mortality," *Journal of Public Economics*, 91(11-12): 2137-2157.

Evans, W.N. & T.J. Moore (2012): "Liquidity, Economic Activity, and Mortality," *Review of Economics and Statistics*, 94(2): 400-418.

Farhi, E. & X. Gabaix (2020): "Optimal Taxation with Behavioral Agents," *American Economic Review*, 110(1): 298-336.

Foley, C.F. (2011): "Welfare Payments and Crime," *Review of Economics and Statistics*, 93(1): 97-112.

Fuchs-Schündeln, N. & T.A. Hassan (2016): "Natural Experiments in Macroeconomics," *Handbook of Macroeconomics*, 2: 923-1012.

Gelman, M., S. Kariv, M.D. Shapiro, D. Silverman, & S. Tadelis (2014): "Harnessing Naturally Occurring Data to Measure the Response of Spending to Income," *Science*, 345(6193): 212-215.

Goldin, J., T. Homonoff, & K. Meckel (2022): "Issuance and Incidence: SNAP Benefit Cycles and Grocery Prices," *American Economic Journal: Economic Policy*, 14(1): 152-178.

Gruber, J. & D.A. Wise (1999): "Introduction to Social Security and Retirement around the World," in *Social Security and Retirement around the World. International Social Security*, eds. J. Gruber and D.A. Wise, Chicago: University of Chicago Press.

Hastings, J. & E. Washington (2010): "The First of the Month Effect: Consumer Behavior and Store Responses," *American Economic Journal: Economic Policy*, 2(2): 142-162.

Hastings, J. & J.M. Shapiro (2018): "How Are SNAP Benefits Spent? Evidence from a Retail Panel," *American Economic Review*, 108(12): 3493-3540.

Huffman, D. & M. Barenstein (2005): "A Monthly Struggle for Self-Control? Hyperbolic Discounting, Mental Accounting, and the Fall in Consumption between Paydays," IZA Discussion Paper.

Kolsrud, J., C. Landais, D. Reck, & J. Spinnewijn (2023): "Retirement Consumption and Pension Design," forthcoming, *American Economic Review*.

Kueng, L. (2015): "Revisiting the Response of Household Spending to the Alaska Permanent Fund Dividend using CE Data," *mimeo*, Swiss Finance Institute.

Kueng, L. (2018): "Excess Sensitivity of High-Income Consumers," *Quarterly Journal of Economics*, 133(4): 1693-1751.

Laibson, D. (1997): "Golden Eggs and Hyperbolic Discounting," *Quarterly Journal of Economics*, 112(2): 443-478.

MacDonald, J.M. (2000): "Demand, Information, and Competition: Why Do Food Prices Fall at Seasonal Demand Peaks?" *The Journal of Industrial Economics*, 48(1): 27-45.

Maloch, F. & C.R. Weaver (1969): "Orientation of Expenditures to Day of Pay," *The Journal of Consumer Affairs*, 3(2): 137-144.

Mastrobuoni, G. & M. Weinberg (2009): "Heterogeneity in Intra-Monthly Consumption Patterns, Self-Control, and Savings at Retirement," *American Economic Journal: Economic Policy*, 1(2): 163-189.

Nakamura, E. & J. Steinsson (2008): "Five Facts About Prices: A Reevaluation of Menu Cost Models: Supplementary Material," *Quarterly Journal of Economics*, 123(4): 1415-1464.

Nevo, A. & K. Hatzitaskos (2006): "Why Does the Average Price Paid Fall during High Demand Periods," *mimeo*, Wharton.

OECD (2021): "Pensions at a Glance 2021: Country Profiles – Japan," https://www.oecd.org/els/public-pensions/PAG2021-country-profile-Japan.pdf.

Olafsson, A. & M. Pagel (2018): "The Liquid Hand-to-Mouth: Evidence from Personal Financial Management Software," *Review of Financial Studies*, 31(11): 4398-4446.

Olafsson, A. & M. Pagel (2020): "Repaying Consumer Debt and Increasing Savings After Retirement," *mimeo*, Columbia Business School.

Parker, J.A. (2017): "Why Don't Households Smooth Consumption? Evidence from a 25 Million Dollar Experiment," *American Economic Journal: Macroeconomics*, 9(4): 153-183.

Parsons, C.A. & E.D. Van Wesep (2013): "The Timing of Pay," *Journal of Financial Economics*, 109(2): 373-397.

Sahm, C.R., M.D. Shapiro, & J. Slemrod (2012): "Check in the Mail or More in the Paycheck: Does the Effectiveness of Fiscal Stimulus Depend on How It Is Delivered?" *American Economic Journal: Economic Policy*, 4(3): 216-250.

Sahm, C.R., M.D. Shapiro, & J. Slemrod (2015): "Balance-sheet Households and Fiscal Stimulus: Lessons from the Payroll Tax Cut and Its Expiration," NBER Working Paper, No. 21220.

Shapiro, J.M. (2005): "Is There a Daily Discount Rate? Evidence from the Food Stamp Nutrition Cycle," *Journal of Public Economics*, 89(2-3): 303-325.

Shimizutani, S. & T. Oshio (2016): "Public Pension Benefits Claiming Behaviour: New Evidence from the Japanese Study on Ageing and Retirement," *The Japanese Economic Review*, 67(3): 235-256.

Shoji, T. (2020): "Liquidity Constraints and a Surge in Storable Goods Demand: Evidence from Japan," *mimeo*, Seikei University.

Stephens Jr., M. (2003): "'3rd of tha Month': Do social Security Recipients Smooth Consumption between Checks," *American Economic Review*, 93(1): 406-422.

Stephens Jr., M. (2006): "Paycheque Receipt and the Timing of Consumption," *The Economic Journal*, 116: 680-701.

Stephens Jr., M. & T. Unayama (2011): "The Consumption Response to Seasonal Income: Evidence from Japanese Public Pension Benefits," *American Economic Journal: Applied Economics*, 3(4): 86-118.

Stephens Jr., M. & T. Unayama (2012): "The Impact of Retirement on Household Consumption in Japan," *Journal of the Japanese and International Economies*, 26(1): 62-83.

Sun, L. & S. Abraham (2021): "Estimating Dynamic Treatment Effects in Event Studies with Heterogeneous Treatment Effects," *Journal of Econometrics*, 225(2): 175-199.

Takayama, N. (2001): "Pension Reform in Japan at the Turn of the Century," *The Geneva Papers on Risk and Insurance*, 26(4): 565-574.

Thaler, R.H. (1999): "Mental Accounting Matters," *Journal of Behavioral Decision Making*, 12(3): 183-206.

Vellekoop, N. (2018): "Explaining Intra-Monthly Consumption Patterns: The Timing of Income or the Timing of Consumption Commitments?" *Leibniz Institute for Financial Research (SAFE)* Working Paper.

Warner, E.J. & R.B. Barsky (1995): "The Timing and Magnitude of Retail Store Markdowns: Evidence from Weekends and Holidays," *Quarterly Journal of Economics*, 110(2): 322-352.

Zeldes, S.P. (1989): "Consumption and Liquidity Constraints: An Empirical Investigation," *Journal of Political Economy*, 97(2): 305-346.

Zhang, C.Y. (2023): "Consumption Responses to Pay Frequency: Evidence from 'Extra' Paychecks," *mimeo*, University of Wisconsin-Madison.

Online Appendix to

Coming in at a Trickle: The Optimal Frequency of Public Benefit Payments

by Cameron LaPoint (Yale SOM) & Shogo Sakabe (Columbia)

Contents

A Model Derivations **51**
A.1 General Expression for the Welfare Loss . . . 51
A.2 Details for Quasi-hyperbolic Discounting Case . . . 52
A.3 Details for Payday Liquidity Case . . . 53
A.4 Optimal Pay Frequency with Consumption Commitments . . . 53
A.5 Menu Costs and Retailer Responses . . . 55
A.6 Counterfactuals with Benefit Eligibility Criteria . . . 59

B Further Details on Expenditure Data **60**
B.1 Constructing Spending Subcategories . . . 61
B.2 Sample Size and Summary Statistics by Trip Frequency . . . 61
B.3 Comparison to Expenditure Diary Survey Data . . . 64

C Using Google Data to Test for Inattention **66**
C.1 Constructing Daily Google Search Data . . . 69
C.2 Estimation Procedure . . . 70

D Temporary Sales around Payday **73**

E Japanese Pension System Branch Offices **78**

F Additional Results on Payday Responses **82**
F.1 Expenditure Responses by Category, Trip Frequency, & Income Decile . . . 82
F.2 Accounting Decomposition of Store-Level Price Changes . . . 86
F.3 Retailer Pricing Responses . . . 87
F.4 Spending Responses around Retirement Age . . . 88
F.5 Spending Responses by Age . . . 93
F.6 Spending Responses by Store Quality . . . 95

A Model Derivations

A.1 General Expression for the Welfare Loss

We can rearrange the compensating variation equation in (2.5) to obtain an intuitive expression for the welfare loss:

$$1 - \lambda(T) = 1 - \frac{T \cdot u^{-1}\left\{ \frac{1}{T}\int_0^T u\Big(c_0 \cdot \exp\Big(-f(t)\Big)\Big)dt \right\}}{\int_0^T c_0 \cdot \exp\Big(-f(t)\Big)dt} \tag{A.1}$$

where $c_0 \equiv \exp(\theta)$ is the value of consumption on payday, as defined in (2.4). Note that in general c_0 is a function of T, but we suppress this dependency for ease of exposition. The numerator in this expression represents the amount of total consumption in a counterfactual scenario where daily consumption is such that the pensioner receives the average daily utility over the actual consumption path. The denominator is total consumption over the pay cycle. Let us denote empirically observed total consumption by C^{tot}. Then we can compactly write the welfare loss as:

$$1 - \lambda(T) = 1 - \frac{T \times \overline{C}}{C^{tot}} = \frac{C^{tot} - T \times \overline{C}}{C^{tot}} \tag{A.2}$$

We can now examine how the welfare loss varies with the government's choice of the interval T between payments. By the quotient rule we can write the marginal compensating variation as

$$\lambda'(T) = \frac{(T \times \overline{C})' \cdot C^{tot} - (T \times \overline{C}) \cdot (C^{tot})'}{(C^{tot})^2} \tag{A.3}$$

where the derivatives of $\overline{C}$ and C^{tot} with respect to T are

$$(\overline{C})' = \frac{\partial}{\partial T} u^{-1}\Bigg\{ \underbrace{\frac{1}{T}\int_0^T u\Big(c_0 \cdot \exp(-f(t))\Big)dt}_{\equiv \overline{U}(T)} \Bigg\}$$

$$= \frac{\partial}{\partial T} u^{-1}\Big(\overline{U}(T)\Big) \cdot \frac{\partial \overline{U}(T)}{\partial T} = u^{-1}\Big(\overline{U}(T)\Big)(u^{-1})'\Big(u(\overline{U}(T))\Big) \cdot \frac{\partial \overline{U}(T)}{\partial T}$$

$$= u^{-1}\Big(\overline{U}(T)\Big)(u^{-1})'\Big(u(\overline{U}(T))\Big) \cdot \frac{1}{T}\left[u\Big(c_0 \cdot \exp(-f(T))\Big) - \frac{1}{T}\int_0^T u\Big(c_0 \cdot \exp(-f(t))\Big)dt \right] \tag{A.4}$$

$$(C^{tot})' = \frac{\partial}{\partial T}\int_0^T c_0 \cdot \exp(-f(t)) = c_0 \cdot \exp(-f(T)) = \overline{B} \tag{A.5}$$

Note that if the Envelope theorem held here, the $\overline{C}'$ term would be zero, since the agent would maintain the same daily average level of utility regardless of the interval length.

Thus, combining (A.3)-(A.5) we can write the FOC to the government's problem in (2.7) as

$$\frac{p(\overline{C} \times C^{tot}) + p(T \times (\overline{C})' \cdot C^{tot}) - p(T \times \overline{C}) \cdot c_0 \cdot \exp(-f(T))}{(C^{tot})^2} = \mu'(T) + p \cdot \overline{B} \tag{A.6}$$

We have now obtained an expression that defines the optimal payment frequency T^* as an implicit function of model parameters and the administrative cost function $\mu(T)$. For each case of our model, we solve (A.1) using standard algorithms for solving nonlinear equations.

A.2 Details for Quasi-hyperbolic Discounting Case

For QH discounters with log utility and $f(t) = \nu \cdot t$, it is straightforward to compute a closed-form expression for the welfare loss:

$$1 - \lambda(T) = 1 - \frac{c_0}{\overline{B}} \cdot \exp\Big(- \nu \cdot T/2\Big) \tag{A.7}$$

For isoelastic utility functions with $\rho \neq 1$, we can define the welfare loss in terms of $\overline{C}$ and C^{tot}:

$$\overline{C} = u^{-1}\left\{ \left(\frac{c_0^{1-\rho}}{\nu \cdot T(1-\rho)^2} \right) \Big[1 - \exp\Big((\rho - 1)\nu \cdot T \Big) \Big] \right\} \tag{A.8}$$

$$C^{tot} = \frac{c_0}{\nu} \cdot \Big(1 - \exp(-\nu \cdot T) \Big) = \overline{B} \cdot T \tag{A.9}$$

Putting everything together, we get the welfare loss for QH discounters with isoelastic utility and inverse IES ρ:

$$1 - \lambda(T) = \begin{cases} 1 - \frac{c_0}{\overline{B}} \cdot \exp\Big(- \frac{\nu \cdot T}{2} \Big) & \text{if } \rho = 1 \\ 1 - \frac{c_0}{\overline{B}} \cdot \left[\frac{1 - \exp\Big((\rho - 1)\nu \cdot T \Big)}{\nu \cdot T(1-\rho)} \right]^{\frac{1}{1-\rho}} & \text{if } \rho \neq 1 \end{cases} \tag{A.10}$$

where, regardless of ρ, payday consumption c_0 is given by

$$c_0 = \frac{\overline{B}(\nu \cdot T)}{\Big(1 - \exp(-\nu \cdot T) \Big)} \tag{A.11}$$

The marginal welfare loss is therefore:

$$-\lambda'(T) = \begin{cases} \frac{c_0 \cdot \nu}{2\overline{B}} \cdot \exp(-\nu \cdot T/2) & \text{if } \rho = 1 \\ \frac{c_0}{\overline{B}\nu(1-\rho)^2 T^2} \left[\frac{\exp(\nu(\rho-1)T)-1}{\nu(\rho-1)T} \right]^{\frac{\rho}{1-\rho}} \frac{\left(\exp(\nu(\rho-1)T)(\nu(\rho-1)T-1)+1 \right)}{1-\frac{c_0}{\overline{B}} \exp(-\nu \cdot T/2)} & \text{if } \rho \neq 1 \end{cases} \tag{A.12}$$

A.3 Details for Payday Liquidity Case

Expression (2.15) shows the welfare loss when consumers exhibit payday liquidity with a spike on payday that is independent of the interval length T. When we allow for the spike to depend on interval length, the welfare loss expression is the same, except we write $x(T)$. For the log utility case, it is straightforward to compute the marginal welfare loss:

$$-\lambda'(T) = \frac{(\bar{c}/\overline{B})(1+x(T))^{1/T-1}}{T^2} \left[\left(1+x(T)\right) \cdot \log\left(1+x(T)\right) - Tx'(T) \right] \tag{A.13}$$

From this equation, we see that irrespective of administrative costs, for this type of payday liquid consumer the government faces a tradeoff between lowering the spike magnitude and decreasing subdivision of the welfare loss when deciding to shorten the pay cycle.

For the case where $\rho \neq 1$, we can use the general formula for the welfare loss in (A.1) to compute the marginal welfare loss. In this context $\overline{C}$ and C^{tot} are given by:

$$\overline{C} = \frac{\bar{c}}{T^{1/(1-\rho)}} \left[\left(1+x(T)\right)^{1-\rho} + (T-1) \right]^{\frac{1}{1-\rho}} \tag{A.14}$$

$$C^{tot} = \overline{B} + \frac{\overline{B} \cdot T}{T + x(T)} = \overline{B} \cdot T \tag{A.15}$$

Using these expressions, we obtain the marginal welfare loss when $\rho \neq 1$:

$$-\lambda'(T) = \frac{\bar{c}}{\overline{B}} \cdot \frac{\left((1+x)^\rho - (1+x) \right) T^{\frac{2-\rho}{\rho-1}} \left((1+x)^{1-\rho} + x - 1 \right)^{\frac{1}{1-\rho}}}{(\rho-1)\left(T(1+x)^\rho - (1+x)^\rho + (1+x) \right)} \tag{A.16}$$

A.4 Optimal Pay Frequency with Consumption Commitments

In the basic version of our framework in Section 2, we take as given that consumers are unaware of their propensity to spend above their average consumption level around payday. That is, they are what the behavioral economics literature has referred to as "naive." In contrast, sophisticated consumers are aware of any internality problem and may exhibit a preference for commitment if, when offered an opportunity to set aside resources for future consumption (e.g. via layaway plans), their willingness to pay for such a commitment device is strictly positive. Bryan, Karlan, & Nelson (2010) provide an overview of commitment device examples and evidence for commitment from field experiments.

We generalize our model to allow for sophisticated consumers who, as of payday, decide to withhold a portion of their payment for future consumption. The strength of this commitment motive grows with pay cycle length T, because, as we show theoretically in Section 2 and in the results on duration elasticities in Table 2, the implied welfare losses from non-smoothing grow larger as T increases. The result is that the government finds it easier to satisfy their optimality condition (2.7) and sets a longer interval T^*, since at least some consumers already partially internalize deviations from the first-best, smooth consumption plan. While we cannot calibrate this version of our model directly using our retail data, we feed in moments from the literature measuring preferences for commitment and compare T^* for the pension system with and without sophisticated consumers.

We simulate the consumption path for the three consumer types and different preference parameters under power utility:

1. A naive consumer who corresponds to the decision problem in Section 2.1 for present-biased consumers. As demonstrated in our calibration results in Section 6.3, marginal welfare losses with respect to T are similar if we simply recast the present-bias problem to that of a consumer who exhibits mental accounting where spending is much larger in the periods on and right after payday. We use c_t^* to denote chosen consumption for the naive agent.

2. A sophisticated consumer who is aware of their internality problem. We use c_t^{**} to denote chosen consumption for the sophisticated agent. We solve for this allocation through backwards induction. This involves first solving for the naive consumer's choice in the penultimate period and then iterating backwards to find the c_0^{**} which optimizes discounted utility conditional on the future naive path of consumption.

3. A sophisticated consumer with access to a commitment device on payday ($t = 0$). We model this commitment device as durable goods stockpiling with linear depreciation over the pay cycle to allow for the possibility that sending more infrequent payments helps individuals smooth consumption through this channel. We then solve for the amount of payday consumption withheld $z_0(T)$ as a function of pay cycle length T. In what follows, we refer to this case as the SC problem, and to solve it we need to solve for the allocations in both of the first two cases. To see this, note that we can write the SC problem as:

$$\max_{z_0} \left\{ u(c_0^{**} - z_0) + \beta \sum_{t=1}^{T-1} \delta^t u(c_t^{**} + z_0/(T-1)) \right\} \text{ s.t. } \begin{cases} z_0 \geq 0 \\ c_0^{**} - z_0 > 0 \end{cases} \tag{A.17}$$

   We require $z_0 \geq 0$. Otherwise, if $z_0 < 0$, commitment would mean taking out a payday loan in period 0 financed through reduced consumption in the future to service the debt. The second constraint says that durables spending on payday cannot be so great that it causes consumption to go to zero or negative.

Finally, we show that one can approximate in continuous time the consumption path of the

sophisticated consumer with commitment via:

$$C(t) = \exp\Big(\theta - f(t) + \zeta(t)\Big) \tag{A.18}$$

where $\zeta(t)$ captures how much consumption "pulls back" towards the optimal smooth path due to commitment. As the simulation shows, $\zeta(t)$ can be non-monotonic in t, but the cumulative pull-back $Z(T) = \int_0^T \zeta(t)dt$ has the property $Z'(T) \geq 0$. This means that as the government decreases the frequency of payments (i.e. $T \uparrow$), unmitigated welfare losses grow through $f(t)$, but commitment is more valuable because withholding consumption in the present allows sophisticated consumers to send more consumption to the future where marginal utility is higher. As before, $\exp(\theta) = c_0$, and we set $f(0) = 0$ and $f(t) = \nu \cdot t$ to match the path of log consumption over the pay cycle for a naive consumer, where ν is the average daily decline in consumption we find in the data.

In general, it is difficult to empirically distinguish between naivete and sophistication, with or without commitments. A well-known result is that for log utility ($\rho = 1$), the consumption paths of naive and sophisticated agents are identical, so there is no preference for commitment in (A.21) and $z_0^* = 0$ (i.e. the slackness constraint binds). Further, the fact that our data on spending originate from grocery stores means that we lack information on purchases of durables and highly storable, but not reusable, goods. Vellekoop (2018) shows that households reduce non-durable spending around mortgage and rent payments, which tend to be scheduled to coincide with income receipt. Zhang (2023) finds that larger but more infrequent payments allows household to save up for lumpy durable goods expenditures like cars or kitchen appliances. This literature on durables and services spending around payday motivates our formulation of the problem via (A.21).

Using the approximation in (A.22), for power utility we can rewrite the welfare loss from non-smoothing under commitment as:

$$1 - \lambda(T) = \begin{cases} 1 - \frac{c_0}{B} \cdot \exp\Big(-\frac{\nu \cdot T}{2}\Big) & \text{if } \rho = 1 \\ 1 - \frac{c_0}{B} \cdot \left[\frac{1 - \exp\Big((\rho-1)(\nu \cdot T - Z(T;\rho))\Big)}{(\nu \cdot T - Z(T;\rho))(1-\rho)}\right]^{\frac{1}{1-\rho}} & \text{if } \rho \neq 1 \end{cases} \tag{A.19}$$

where $Z(T;\rho) = 0$ under log utility due to no preference for commitment. We write $Z(T;\rho)$ to emphasize that the pull back to the smooth path depends on preferences for commitment, which in turn depend on the consumer's intertemporal elasticity of substitution.

A.5 Menu Costs and Retailer Responses

We now incorporate monopolistic retailers into the optimal frequency framework. Retailers set prices to maximize real profits, subject to menu costs that must be paid for increasing prices on payday. This puts further downward pressure on the optimal payment frequency, as the government can render the gains from price discrimination small relative to the menu cost by shortening the pay cycle length. We consider two variants of this model. One in

which there is a single monopolistic retailer who sets the price for the entire consumption basket, and another version in which retailers specialize in a particular variety. The intuition is largely the same across the two versions.[1]

A.5.1 A Single Monopolistic Firm

There is a single monopolistic retailer who sets prices for a common basket of goods purchased by all consumers. The retailer maximizes pay cycle profits according to the problem:

$$\max_{\{P_t\}} \left\{ \sum_{t=0}^{T-1} P_t \cdot Y_t - W_t \cdot L_t - \kappa \cdot W_t \times \mathbb{1}_t - P_t \cdot \Gamma \right\} \tag{A.20}$$

where $\kappa \cdot W_t$ represents the real menu cost denoted in κ units of labor L_t, and $\mathbb{1}_t$ is an indicator for whether the retailer changes the price P_t at time t. We model menu costs as a labor cost, in line with observations from firm-level surveys conducted by Blinder et al. (1998). Firms use labor L_t provided by the working population, which in this setting is the non-pensioners or non-recipients (indexed hereafter by "NR" of benefits). Labor is used to satisfy non-recipient demand $F(L_t) = C_t^{NR}$, and the retailer pays a fixed real inventory cost Γ each period to external suppliers who stock their shelves and inelastically satisfy the residual demand of the recipients (C_t^R).[2]. The pay cycle runs from payday at $t = 0$ until $T - 1$, where T is the number of days in between payments. We set the discount rate on future profits to unity for ease of exposition.

We assume claims to retailer profits are held by external (e.g. foreign) shareholders. If we assume all profits were rebated to the working population, then the problem would collapse to the original one outlined in Section 2, since any price discrimination would ultimately be to the benefit of the working population, and then subsequently taxed away by the government to finance the benefits system. Alternatively, if profits are proportionally split between recipients and non-recipients, then whether there is an additional welfare loss depends on how much of the surplus profits accrued to the non-recipients gets eroded by real wage deflation from payday price hikes. While interesting to consider, profit redistribution significantly complicates the model without changing the basic tradeoffs faced by the government.

The fraction p of the population who are non-recipients choose consumption and labor over the pay cycle to solve the following problem:

$$\max_{\{C_t, L_t\}} \left\{ \sum_{t=0}^{T-1} u(C_t) - \nu(L_t) \right\} \quad \text{s.t.} \quad P_t \cdot C_t = S_t + W_t \cdot L_t - \frac{\tau(b)}{T} \tag{A.21}$$

[1]Note in focusing on payday liquidity we adhere to the discrete time representation of the model used in Section 2.3 in the main text.

[2]This business structure is akin to a retailer who has their own private label brand of products sold alongside products received from a third-party wholesaler which they price at a markup. If instead we simply assumed $F(L_t) = C_t$, the retailer would always find it profitable to raise prices on payday, meaning the government could influence price changes along the intensive margin instead of the extensive margin.

Suppose the fraction $(1-p)$ of the population who are pensioners – or more generally, those consumers receiving regular benefits – have payday liquidity. That is, as described in equations (2.12) and (2.13) in Section 2.3, they smooth real consumption throughout the pay cycle at a level $\bar{c}$, but splurge by consuming $\Big(1+x(T)\Big)\cdot\bar{c}$ on payday $(t=0)$.[3]

The government runs a balanced budget as in (2.1) and finances benefits through lump-sum taxation of the non-recipients. Because the government's choice of interval length influences the retailer's pricing strategy, relative to equation (2.6) there is an additional term in the objective function which accounts for the fact that if retailers choose to price discriminate, benefit non-recipients experience a reduction in their maximized level of utility $U^*(\cdot)$:

$$\min_T \Big\{ -p\cdot\lambda(T)+\gamma\cdot\Big(p\cdot b(T)+\mu(T)\Big)+\Big(U^*(C^{1,NR}(T))-U^*(C^{0,NR}(T))\Big)\Big\} \tag{A.22}$$

As in the model without menu costs, the first term represents the purely paternalistic motivation of the government to set a lower T to mitigate the non-smoothing behavior of the payment recipients, and the second term captures the costs of imposing taxes on the working population, with marginal cost of funds $\gamma=1$ in the case of lump-sum taxes. The final term is the gap between the maximized level of utility when the non-recipient's consumption is either altered by payday pricing $\mathbb{1}_0=1$ or smooth for $\mathbb{1}_0=1$.

If the government sets T such that the retailers do not engage in a temporary price hike, then $C^{0,NR}(T)=C^{1,NR}(T)$ and the non-recipients experience no welfare loss relative to the original problem without menu costs and monopolistic pricing. On the other hand, if administrative costs are sufficiently convex, it may make sense for the government to allow price discrimination by setting T such that excess payday demand attracts payday price hikes and generates a loss for the non-recipients.

In any case, this problem depends on a number of additional parameters not discussed in the main text, such as the disutility of labor $\nu(L)$, firm production $F(L)$, and the size of menu costs κ. To illustrate the solution to this problem via closed-form expressions, we now consider a very simple parameterization where retailers have linear production $F(L)=A\cdot L_t$, with a constant productivity A, and workers have log utility from consumption with constant marginal disutility of labor: $u(C_t)-\nu(L_t)=\log(C_t)-\omega L_t$.

First, under this parameterization, we can combine the non-recipient's FOC for consumption and labor supply to show that the real wage is proportional to consumption:

$$\frac{1}{C_t^{NR}}=P_t\cdot\eta_t \tag{A.23}$$

$$W_t\cdot\eta_t=\omega \tag{A.24}$$

$$\implies \frac{W_t}{P_t}=\omega\cdot C_t^{NR} \tag{A.25}$$

[3]The qualitative results of this model obtain for quasi-hyperbolic discounters as well, since the welfare loss from non-smoothing consumption will balloon with respect to the interval length T. Focusing on the payday liquidity case helps illustrate how retailer price discrimination affects the optimal frequency with closed-form expressions for prices and consumption.

where η_t is the Lagrange multiplier on the non-recipient's budget constraint. Using this fact and the linear production function, we can then rewrite the retailer's real profits over the pay cycle as

$$\sum_{t=0}^{T-1} C_t - \omega C_t^{NR} \cdot \frac{C_t^{NR}}{A} - \kappa \cdot \omega C_t^{NR} \times \mathbb{1}_t - \Gamma \tag{A.26}$$

where C_t is aggregate demand for the retailer's products, which is a weighted average of the demands of the non-recipients (NR) and recipients (R):

$$C_t = p \cdot C_t^R + (1-p) \cdot C_t^{NR} \tag{A.27}$$

Since the payday liquid consumers are price-insensitive, in that they pick consumption based on a heuristic rule of thumb, the firm's problem reduces to picking a point on the non-recipient's demand curve to maximize real profits. The firm's FOC with respect to C_t^{NR} yields equilibrium real expenditures of benefit non-recipients:

$$C_t^{NR} = \begin{cases} \frac{A}{2\omega} & \text{if } \mathbb{1}_t = 0 \\ \frac{A\left((1-p)-\kappa\cdot\omega\right)}{2\omega\cdot(1-p)} & \text{if } \mathbb{1}_t = 1 \end{cases} \tag{A.28}$$

Hence, non-recipients smooth consumption unless there is a temporary price hike on payday ($t = 0$), and they consume more in real terms in the absence of a price hike ($\mathbb{1}_t = 0$).

Combining non-recipient expenditures with the condition in equation (A.25) yields the pricing strategy as a function of the nominal wage:

$$P_t = \begin{cases} \frac{2W_t}{A} \text{ if } \mathbb{1}_t = 0 \\ \frac{2W_t\cdot(1-p)}{A\cdot(1-p-\kappa\omega)} \text{ if } \mathbb{1}_t = 1 \end{cases} \tag{A.29}$$

where for a given nominal wage W_t, prices are always higher when $\mathbb{1}_t = 1$, meaning retailers raise prices rather than engage in temporary sales.

Because recipients maintain a smooth consumption profile beyond payday, the retailer will only find it profitable to pay the menu cost to change prices at $t = 0$. This means retail price discrimination, if it does occur, will be concentrated on payday, and the relevant comparison from the retailer's perspective is profits at time zero when $\mathbb{1}_0 = 0$ vs. profits when $\mathbb{1}_0 = 1$. Denote real profits on payday with the price hike as $\Pi_0^{R,1}$, and real profits on payday without the price hike as $\Pi_0^{R,0}$. Then the retailer will find it worthwhile to pay the menu cost to capture some of the excess demand from the recipients on payday whenever the following

condition is met:[4]

$$x(T) > \frac{1-p}{p\bar{c}\cdot\Big((2-A)(1-p)-A\cdot\kappa\omega)\Big)}\cdot\left\{A\kappa+\frac{A(1-p-\kappa\omega)}{(1-p)}\cdot\left(p\bar{c}+\frac{A(1-p)-\kappa\omega}{2\omega}\right)\right\} \tag{A.30}$$

This condition says that the retailers will find it profitable to price discriminate whenever excess demand from the benefit recipients, captured by $x(T)$, is sufficiently high that the increase in profits offsets reduced demand from the non-recipients plus the menu cost.

Finally, the difference in the non-recipients' optimized level of utility when there is price discrimination is given by:

$$\log(C_0^{NR,1})-\log(C_0^{NR,0})+\omega\cdot\Big(L_0^1-L_0^0\Big)=0 \tag{A.31}$$

The difference in the optimized level of utility is ultimately zero in this special version of the model where all price discrimination occurs on the extensive margin. That is because the loss in utility over consumption associated with the price hike is completely offset by the reduction in disutility from labor supply. Therefore there exist parameter sets (p, A, κ, ω) such that even if the government optimally chooses the pay cycle length to promote consumption smoothing among the benefit recipients, retailers may respond by raising prices on payday.

A.6 Counterfactuals with Benefit Eligibility Criteria

In Section 6.4, we provide estimate the cost savings resulting from a counterfactual normal retirement age (NRA) of 70. To do so, we calculate the implied increase in penalty rates imposed on early pensioners aged 60 to 69 compared to the current NRA of 65 in Japan. Currently, if individuals born on or before April 1, 1964 claim pension payments before their 65th birthday month, they incur a penalty rate of 0.5% per month until reaching the month of their 65th birthday. For example, if a pensioner claimed benefits on the month they turned 60, the penalty rate would be 30%, the maximum under the current NRA. This means the pensioner would receive 70% of the normal pension benefits she would have received had she claimed benefits at age 65, for the remainder of her life.

If the NRA were changed to 70, the penalty rate could increase to 60% under the current rate structure. From the government's perspective, this increased penalty would translate into cost savings resulting from the NRA change. In the analysis below, we also consider scenarios where the penalty rate is capped at the current maximum rate of 30% and where the penalty rate is set at 0.4% per month for individuals born on or after April 2, 1962.

[4]The common component of profits across the two pricing strategies over $1 \leq t \leq T-1$ is:

$$\frac{T-1}{2}\cdot\Big[p\cdot\bar{c}+(1-p)\cdot A/2\omega\Big]$$

We use data from the Statistics Bureau on Employees' Pension Insurance and National Pension Programs.[5] These data include information on the number of pensioners and their average monthly pension payments received broken down by age group. We focus on pensioners eligible solely for the lump-sum portion of the national pension, specifically those between 60 and 69 years old, as 60 is the earliest age one can claim these payments. We exclude late retirees aged 70 and older, which make up around 0.5% of pensioners, as the data does not provide a detailed age distribution for this group. To determine the amount of normal, unpenalized payment amounts these pensioners would receive if they retired neither early nor late, we refer to the penalty and reward rate tables provided by JPS.[6] After calculating the aggregate amount of regular pension payments, we apply the new penalty rates based on the new NRA.

We first consider a scenario in which the penalty rate is 0.5% per month between the time early pensioners claim payments and the month of their 70th birthday. The additional penalties imposed on early pensioners amount to 36.12 billion JPY annually, with 33.06 billion JPY accruing to those aged 60-64 and 3.06 billion JPY to those aged 65-69. The total amount imposed on late pensioners (66 and above) under the original NRA is relatively small, as they account for only 1.8% of pensioners who started receiving payments in 2011.

It is important to note that these estimates solely pertain to the impact on the transitional cohort of pensioners who claimed their pension benefits in 2021. Assuming that all pensioners who retired in 2021 live for an additional 20 years until 2041, total undiscounted savings would reach 722.4 billion JPY (36.12 billion JPY $\times$ 20 years). This comparison enables us to analyze the dynamics of administrative cost savings in relation to accumulated pension cost savings through penalties. In order to estimate the additional effects on subsequent cohorts, we would need to make additional assumptions about demographic transitions in Japan.

If the maximum penalty rate remains capped at 30% – the current maximum rate – the additional penalties in this case amount to 13.26 billion JPY, with 10.2 billion JPY accruing to those aged 60-64 and 3.06 billion JPY to those aged 65-69. Alternatively, if the penalty rate is 0.4% per month, the additional penalties in this case total 28.07 billion JPY, with 25.26 billion JPY accruing to those aged 60-64 and 2.81 billion JPY to those aged 65-69.

B Further Details on Expenditure Data

We provide in this appendix details on how we sort barcodes into spending categories, construct our shopper panel, the representativeness of our shoppers and the fraction of total spending our dataset covers based on alternative survey data sources.

[5] https://www.e-stat.go.jp/stat-search/files?page=1&toukei=00450463&tstat=000001064713

[6] https://www.nenkin.go.jp/service/jukyu/roureinenkin/kuriage-kurisage/index.html

B.1 Constructing Spending Subcategories

Our data contain 968 major goods categories provided by Magee Co., with some categories containing thousands of barcodes. The majority of these categories consist of perishable food items. However, while most of the goods in our data would be considered non-durable by standard definitions proposed in the empirical literature on consumption patterns, we find it useful to further aggregate these categories based on whether they consist of fresh foods or packaged foods, alcohol, tobacco, or other non-discretionary spending. Relative to total expenditures, these subcategories provide information on how diet composition varies over the pay cycle, and information on which goods are more likely to be purchased on paydays.

In the end, we consider 13 subcategories, plus a category called "raw foods" which contains all fresh, non-packaged food items in our sample. We summarize the contents of these subcategories in Table B.1. Table B.2 and Table B.3 provide summary statistics for each subcategory for our preferred panel of weekly shoppers buying perishables. Further summary of the goods included in these subcategories can be found in the Online Data Appendix. Some stores drop out of our sample when we focus on less common goods categories such as fresh fish, tobacco, and "other," which includes mostly non-food items available at the cash register (e.g. gift cards).

B.2 Sample Size and Summary Statistics by Trip Frequency

We face an inherent tradeoff in defining our estimation sample of shoppers in the retail scanner data. Our goal is to capture the consumption profile at daily frequency $C(t)$ for the typical individual in the economy. Shoppers who regularly go to the store in our dataset may or may not provide a more accurate snapshot of $C(t)$; while such shoppers are less likely to stockpile goods, they are likely not representative of the population as a whole. For instance, frequent shoppers might have preferences skewed in favor of prepared meals. As Figure B.1 indicates, we face a bias-variance tradeoff in that as we restrict attention to more frequent shoppers, our sample size diminishes. There are roughly 4 million shoppers with loyalty point cards in our uncensored data, but this number declines by 90% once we focus on weekly shoppers ($k \geq 4$).

To assess the direction of biases in our measures of $C(t)$ from subsetting to more vs. less frequent shoppers, we plot in Figure B.2 and Figure B.3 daily spending broken down by goods category and age group as a function of the shopper's minimum number of visits per month k. Several interesting patterns emerge. First, in Figure B.2 overall spending peaks at $k = 3$ and then declines almost linearly before trending upward at $k > 12$. For shoppers who visit the store less than 3 times per week on average, the daily consumption profile is similar regardless of whether the shopper is above or below the normal retirement age (65 years old).[7] The relationship between visit frequency and daily spending differs dramatically

[7]These trends are partially driven by the unbalanced nature of the panel. Shoppers who were early adopters of the point card will have a longer panel, or a larger denominator used to compute average $C(t)$. This dimension of selection is not due to the retailer's decision to use the market research firm because we restrict to a balanced panel of stores.

Table B.1. One-digit Goods Subcategory Contents

One-digit Category	Two-digit Category	Four-digit Categories
Fresh fruits & vegetables	Fresh fruits	seasonal fruits, imported fruits, assorted fruits, fruit-related products,
	Fresh vegetables	leafy veg., stalk veg., root crops, edible plants, edible seeds, mushrooms, germinated veg., assorted veg.
Processed fruits & vegetables	Processed fruits	frozen fruits, cut fruits,
	Processed vegetables	boiled veg., frozen veg., cut veg.
Fresh fish	Fresh fish	round items, filet, shellfish, assorted fish
	Sashimi	brick form, sashimi, tataki, raw fish, assorted fresh fish
Preserved fish products	Salted & dried fish	boiled fish, frozen fish, seasoned fish, pickled fish, salted fish, dried fish, fish eggs, seaweed
Raw meat & poultry	Beef	wagyu, domestic beef, imported beef,
	Pork	domestic pork, imported pork
	Chicken	domestic chicken, imported chicken, brand name chicken, duck meat
	Meat varieties	lamb, horse meat, minced meat, offal, raw meat, eggs, dairy products
Grains	Cereals	powder, rice, mochi, raw noodles, dough, bread, cereal
Other processed foods	Seasonings	cooking oil, spices, condiments, spread/dips, toppings, rice seasoning
	Dry produce	dried fish, dried fruits
	Processed food	pickled items, processed fish, pastes, cooked beans, processed meats
	Instant foods	cup noodle, instant soup, frozen foods, sealed rice pouch
Prepared foods	Semi-prepared dishes	fried, simmered, grilled, Japanese, Western, Chinese
	Side dishes	fried, grilled, grilled eel, Japanese, Western, Chinese
	Bento	cooked rice, sushi, bread dishes, noodle dishes
Sweets and desserts	Confectionery	toppings, jelly/pudding, ice cream, frozen confections, candies/cookies, rice crackers
Non-alcoholic beverages	Beverages	coffee/tea, milk-based drinks, vegetable/fruit drinks, soft drinks
Alcohol	Alcohol	beer, liqueurs, wine liquor, sake
Tobacco	Tobacco	tobacco
Other discretionary	Other	flowers, gifts/confections, kiosk goods, service counter goods

Table B.2. Summary Statistics by Goods Subcategory Expenditures

A. Fresh Food Categories

	Fruits/Vegetables	Fish	Meat/Poultry	Prepared Foods
Avg. daily expenditures (JPY)	519	581	605	545
Avg. monthly expenditures (JPY)	5,635	4,270	5,565	5,330
Avg. number of monthly trips	9.9	6.7	8.3	8.5
% Normal retirement age	45.6%	54.5%	42.7%	45.9%
% Early retirement age	31.3%	37.4%	28.7%	31.6%
# Stores	510	494	508	508
# Shoppers	391,378	128,629	341,799	243,930

B. Processed Food Categories

	Fruits/Vegetables	Fish	Grains	Other Processed
Avg. daily expenditures (JPY)	222	397	399	662
Avg. monthly expenditures (JPY)	1,149	2,501	3,382	7,455
Avg. number of monthly trips	4.6	5.9	7.9	10.4
% Normal retirement age	37.1%	52.8%	42.4%	45.3%
% Early retirement age	25.3%	36.1%	28.5%	31.0%
# Stores	503	504	509	510
# Shoppers	22,825	101,954	271,449	401,640

C. Discretionary Goods Categories

	Sweets/Desserts	Non-alcoholic Beverages	Alcohol	Tobacco	Other
Avg. daily expenditures (JPY)	374	370	847	1,080	479
Avg. monthly expenditures (JPY)	3,167	3,302	6,796	7.739	2,789
Avg. number of monthly trips	7.9	8.1	7.7	6.0	5.7
% Normal retirement age	40.5%	40.3%	34.2%	39.1%	48.9%
% Early retirement age	27.1%	27.1%	21.7%	25.1%	34.2%
# Stores	510	509	508	381	474
# Shoppers	237,050	260,936	50,186	5,333	22,474

Table B.3. Aggregate Retail Expenditures Summary Statistics

	All Goods	Raw Foods (Perishables)
Avg. daily expenditures (JPY)	2,603	1,149
Avg. monthly expenditures (JPY)	35,184	14,048
Avg. number of monthly trips	13.0	11.5
Avg. periodicity	2.0	2.3
% Female shopper	65.7%	65.5%
% Early retirement age	45.5%	45.5%
% Normal retirement age	31.2%	31.5%
# Stores	511	510
# Shoppers	409,439	416,726

Notes: The table reports summary statistics for purchases on all goods and purchases on raw foods, which include all perishable food items. Monthly and per-trip expenditures are reported in nominal Japanese yen. Periodicity refers to the number of days between consecutive shopping trips. % Female shopper refers to whether the shopper who initially signed up for the point card reported female as their gender. % Early retirement age refers to the percentage of shoppers who had reached the earliest possible age (60 years old) for claiming benefits as of the beginning of the panel, and % Normal retirement age refers to the percentage of shoppers who attained the normal retirement age (65 years old). All statistics were computed from a sample of shoppers aged 20 to 90 years old who visited a store in our sample to buy perishables at least four times per month.

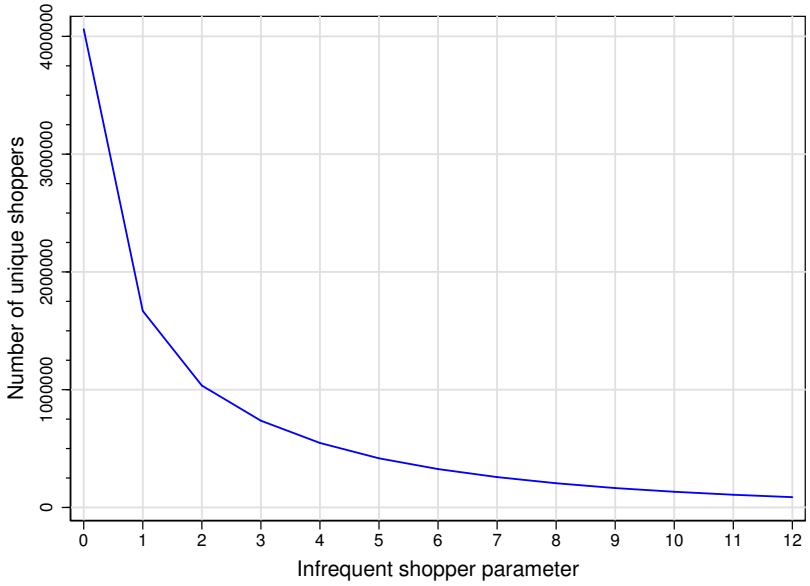
across the type of good. The slope is positive with respect to visit frequency for the most perishable goods like fresh fish, meat/poultry, and prepared foods, but negative or relatively non-monotonic for goods which have longer shelf lives like sweets/desserts, alcohol, and grains. Overall, by using weekly shoppers in our baseline analysis, we overestimate average daily perishables consumption by, at most, 9%, but this decision leads us to underestimate daily consumption of certain goods in a way that appears unrelated to the magnitude of the spike in payday spending.

B.3 Comparison to Expenditure Diary Survey Data

As noted in the main text, one concern with using retail scanner data to measure pay cycle consumption is that we may not fully capture the actual consumption profile, because shoppers may go in and out of our sample due to switching away from stores contained within the data provider's network and/or failing to use their loyalty point card to log purchases. We consult the Japanese Family Income and Expenditure Survey (FIES) to benchmark spending measured in our retail data to average monthly spending over a nationally representative household survey.[8]

[8]The FIES data files we use for this exercise are publicly available here: https://www.e-stat.go.jp/stat-search/files?page=1&toukei=00200561&tstat=000000330001. See the Online Data Appendix for tabulations from the FIES and details on how we map spending categories between the FIES and retail data.

FIGURE B.1. Regular Shoppers with $\geq k$ Store Visits Each Month



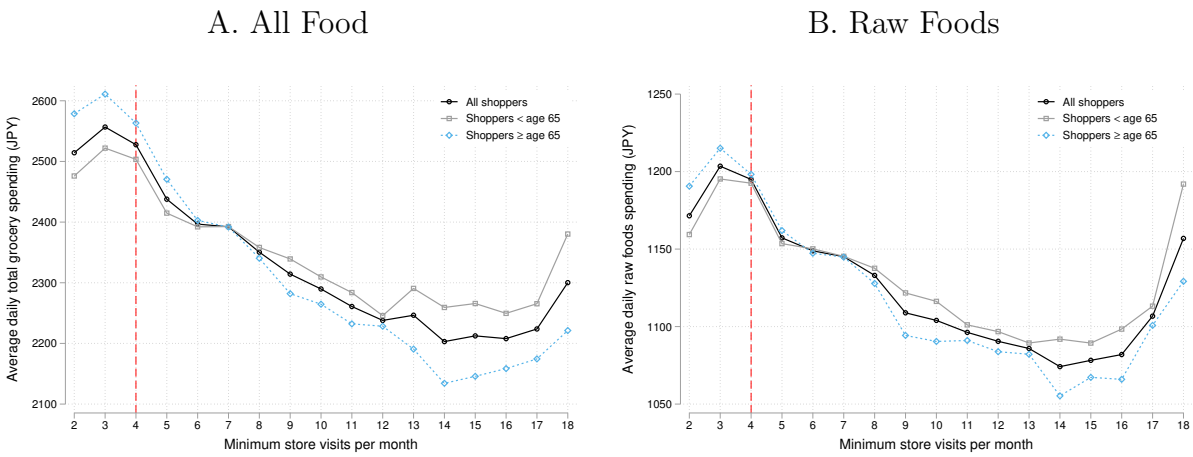
Notes: The figure shows the number of unique shoppers, identified by loyalty point card number, who make at least k visits (x-axis) to a store in our sample during each month of our sample period. $k = 0$ corresponds to the entire sample of shoppers, which includes all shoppers who go to the store and use their point card at least once over the entire period April 2011 – October 2014.

Figure B.4 plots how average monthly spending varies with the monthly store visit frequency parameter k, against the corresponding FIES average within each category displayed as a dashed red line. Although our baseline estimation sample of weekly shoppers ($k \geq 4$) results in average monthly total grocery spending at 50% of equivalent measures in the FIES, we can match average FIES spending for very frequent shoppers who visit a store every other day, on average ($k \geq 16$). We isolate this sample of very frequent shoppers in several robustness checks in Appendix F.1 and Figure 3 in which we show that our results are qualitatively and quantitatively unchanged if we use a sample of shoppers for which we can match average monthly spending in the FIES. The cutoff frequency parameter k needed to match FIES spending for respondents above the early retirement age (60 years old) is 17, or 15 for respondents below age 60.

However, the FIES is compiled with measurement error, and even at low trip frequencies the retail data show much higher average monthly spending on certain types of items such as alcohol and pantry goods used for cooking ("other processed foods" in Table B.1).[9] For instance, average monthly purchases on alcohol total 7,028 JPY for weekly shoppers in the retail data, but only 2,726 JPY in the FIES; for pantry goods the numbers are 5,925 vs. 2,631, respectively. Unayama (2018) compares tabulations across different Japanese spending

[9]Note that we cannot directly compare the public FIES spending totals for separate fresh food categories like fruits or vegetables, but we can compare perishables (our "raw foods" categorization) across the two data sources. To map to our retail data, we define "all food" in the FIES as all food spending excluding the subcategories: "meals outside the home" and "board," which includes kitchen supplies.

FIGURE B.2. Average Daily Spending for All Food and Raw Foods



Notes: We plot average daily grocery spending for all food spending and perishables ("raw foods") as a function of the minimum number of visits (k) a shopper makes to the store each month. We do this separately for shoppers above or below the normal retirement age (65 years old), as well as for the entire sample. To limit computational run time, we restrict to $k \geq 2$ – that is, shoppers who visit the store at least two times each month while in sample. 18 trips each month is the 99th percentile of k in our sample. A shopper's daily spending on a given date is coded as zero if we do not have record of that shopper visiting a store on that date. The vertical red dashed line at $k = 4$ in each figure panel indicates average daily spending for our preferred sample of shoppers who, on average, visit the store on a weekly basis.

surveys to show that respondents systematically omit lumpy expenditures like cars and kitchen appliances in their FIES spending diaries; it is well-documented that durables and storeable goods are under-reported in the CEX as well (Aguiar & Bils 2015).

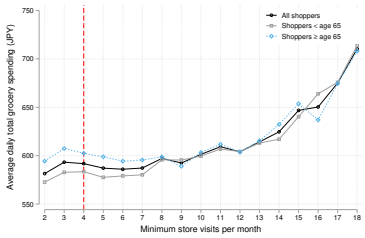
The average share of food in overall monthly household spending is 24%, or 25% for those above normal retirement age. The share of perishables in grocery spending is 45.5% in the FIES vs. 45.6% for raw foods spending of weekly shoppers in the scanner data. Hence, the expenditure weight on instantaneous consumption is roughly constant regardless of store visit frequency.

C Using Google Data to Test for Inattention

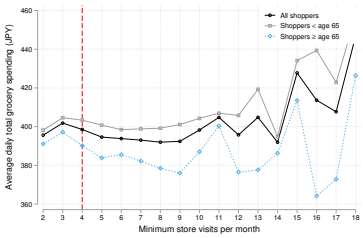
This appendix outlines how we construct daily time series of Google searches related to payments from the Japanese National Pension System. We use these series to gauge the extent to which the timing of upticks in search activity line up with scheduled payment dates. This exercise is useful for two reasons: (i) it allows us to assess whether shoppers anticipate pension payments, and (ii) it allows us to discern whether shoppers are inattentive to changes in the payment schedule due to holidays and weekends. Baugh & Wang (2021) argue that consumers may be unaware of or ignore rules governing regular Social Security payment schedules, even when such rules are public information. Contrary to this notion, we find a clear uptick in searches on the eve of a scheduled payment date.

FIGURE B.3. Average Daily Spending by Subcategory and Trip Frequency

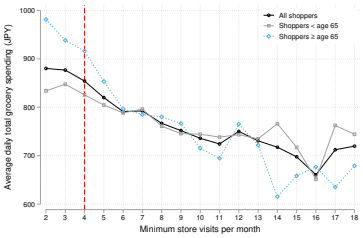
(a) Prepared Foods



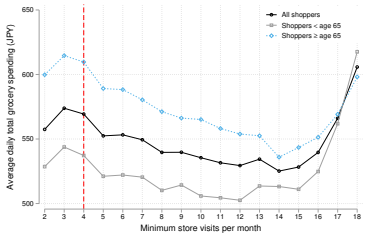
(b) Sweets/Desserts



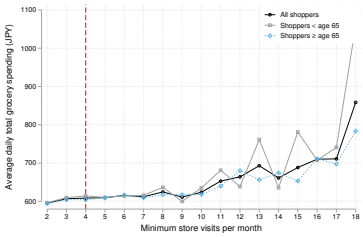
(c) Alcohol



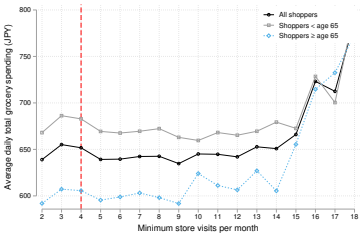
(c) Fresh Produce



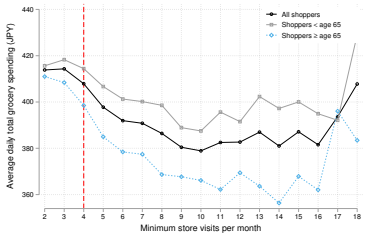
(d) Fresh Fish



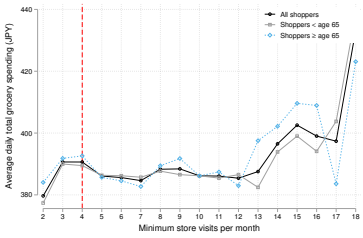
(e) Meat & Poultry



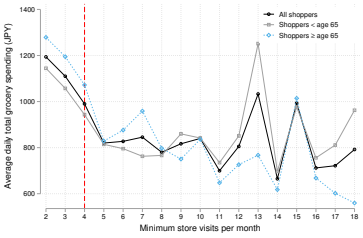
(f) Grains



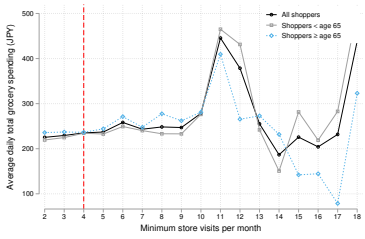
(g) Non-alcoholic beverages



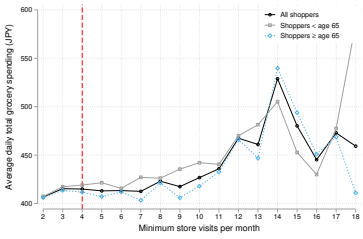
(h) Tobacco



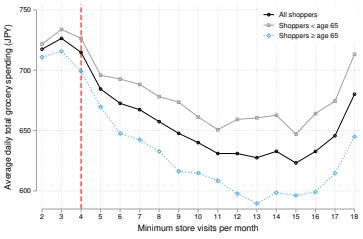
(i) Processed Fruits/Vegetables



(j) Preserved Fish

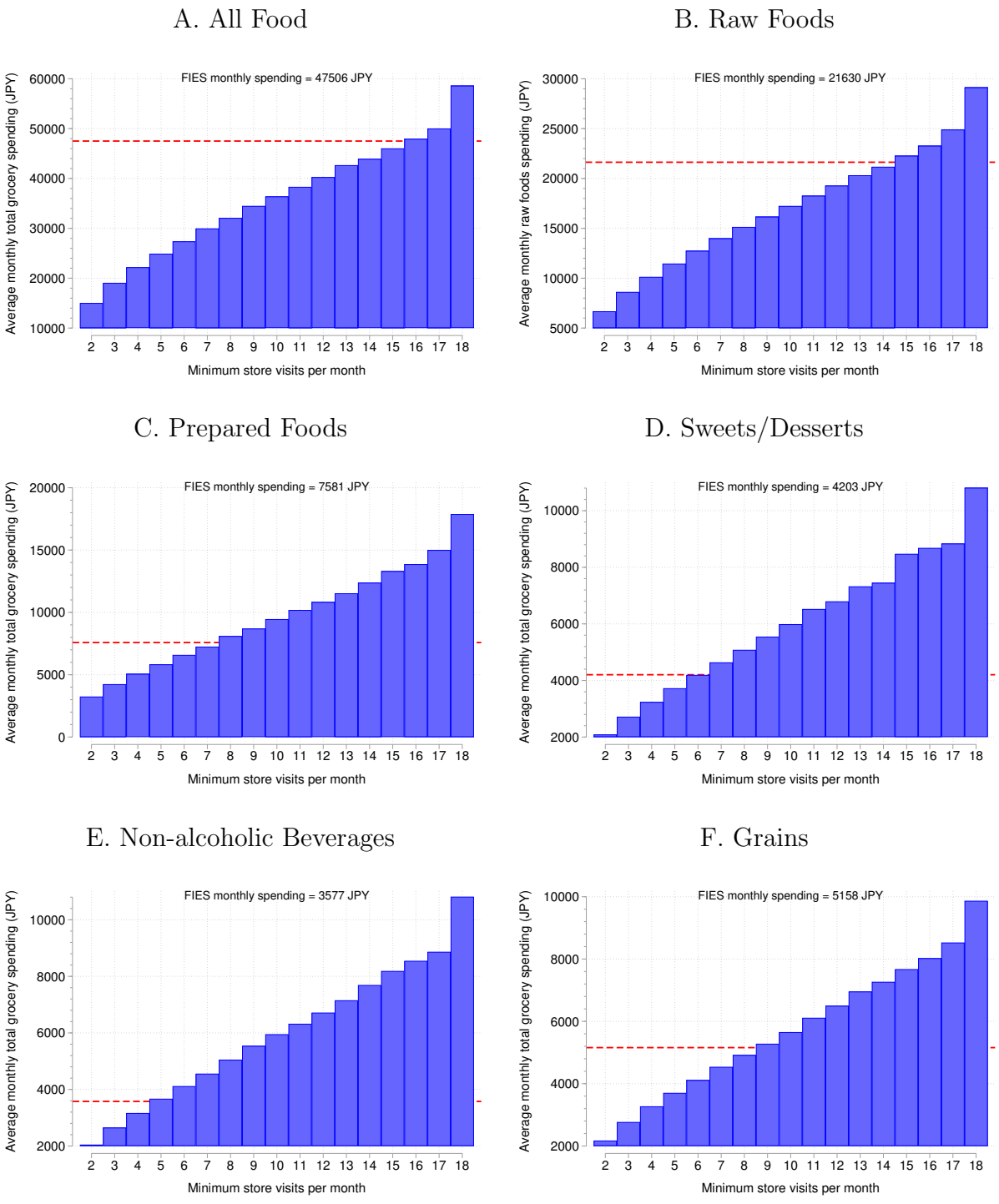


(k) Other Processed Foods



Notes: We plot average daily grocery spending for major goods subcategories as a function of the minimum number of visits (k) a shopper makes to the store each month. We do this separately for shoppers above or below the normal retirement age (65 years old), as well as for the entire sample. To limit computational run time, we restrict to $k \geq 2$ – that is, shoppers who visit the store at least two times each month while in sample. 18 trips each month is the 99th percentile of k in our sample. A shopper's daily spending on a given date is coded as zero if we do not have record of that shopper visiting a store on that date. The vertical red dashed line at $k = 4$ in each figure panel indicates average daily spending for our preferred sample of shoppers who, on average, visit the store on a weekly basis. See Table B.1 for precise goods category definitions.

FIGURE B.4. Average Monthly Spending: FIES vs. Scanner Data Shoppers



Notes: The figure compares average monthly spending by shoppers' store visit frequency to the same statistic tabulated from the consumption diaries produced as part of the Japanese Family Income and Expenditure Survey (FIES). The red dashed line in each figure panel corresponds to average monthly spending in the FIES during the years of our sample time period 2011–2014. We compare total food spending, raw foods (perishable goods spending), and four subcategories of grocery spending in the public FIES data files which directly map to 1-digit categories in our scanner data. To limit computational run time, we restrict to $k \geq 2$ – that is, shoppers who visit the store at least two times each month while in sample. We lump into a single bucket shoppers who make more than 18 trips to the store each month, which is the 99th percentile of k in our sample. See text for more information on the FIES data and Table B.1 for goods category definitions.

C.1 Constructing Daily Google Search Data

When querying Google Trends, the user provides search terms, a region code (e.g. "JP" for Japan), and a time range. Google Trends data are available from January 1, 2004.[10] It is possible to add punctuation to queries to filter the results. For instance, one can wrap double quotation marks around keywords to restrict to searches containing the exact phrase. One can also use the addition operator to generate intersected queries. In our application, we use the minus operator to search for "public pension payments - fees," which excludes searches related to contributions from participants who have not yet begun to claim benefits.[11]

Instead of reporting the total amount of searches for the user-provided keywords, Trends reports a search volume index (SVI). For example, if one typed "pension" into the Google Trends search bar, for each t in time range τ and region r one would then obtain:

$$\text{pension search index}_{r,t} = 100 \times \frac{\text{SP}_{r,t}}{\max_r \text{SP}_{r,t}} \tag{C.1}$$

where the search propensity (SP) is computed as:

$$\text{SP}_{r,t} = \frac{\Big[\text{Google searches containing "pension"}\Big]_{r,t}}{\Big[\text{total Google searches}\Big]_{r,t}} \tag{C.2}$$

The SVI is a function of the aggregate search volume Google receives. This raises the issue that some of what the SVI measure captures is trends in total Google searches over time. We check that this denominator effect is not confounding our results using total page views from the Wikipedia page for the Japanese public pension system, rescaling the page views series as Google does, and then comparing results using the scaled vs. unscaled page views.

Google Trends, by default, returns a monthly series for time ranges longer than 5 years, a weekly series for time ranges shorter than 5 years but longer than 8 months, and a daily series for time ranges spanning less than 8 months. Hence, to obtain SVI data at the daily frequency for a long time period, the user must generate separate queries for several subperiods. Since the index level is relative to the week with the greatest search volume within that subperiod, the user must create a common index across the subperiods to maintain the interpretation of the SVI as a measure of relative search propensity over the entire sample period.

Therefore to create a weekly series we perform the following steps:

1. For each keyword search, download the full monthly series for 2004-2018.
2. Divide the period into subperiods τ_k of equal length and download the weekly data for each subperiod. Merge the subperiods together into a single data frame.

[10] Note that Google frequently changes the Trends interface, so some procedures discussed here may not be applicable in the future. The current web browser version of the tool can be accessed at http://trends.google.com/trends/explore

[11] Our results are qualitatively similar for non-interacted keyword searches of "public pension payments."

3. Create a weekly data frame that assigns to each week the SVI observation for the corresponding month.

4. Create another weekly data frame that assigns to each week j the SVI observation in the first week j' of the corresponding month i.

5. Rescale the downloaded weekly SVI series via:

$$Z^*_{i,j} \equiv \underbrace{\frac{\text{SVI}_{i,\tau}}{\text{SVI}_{j',\tau_k}}}_{\text{correction factor}} \times \underbrace{\text{SVI}_{i,j,\tau_k}}_{\text{downloaded data}} \tag{C.3}$$

   For cases where the first week of month i has an SVI of zero, normalize $Z^*_{i,j} = 0$.

6. Create the new series $SVI^*_{i,j,\tau}$ from the data frame of scale factors Z^*:

$$SVI^*_{i,j,\tau} \equiv \text{round}\left(\frac{Z^*_{i,j}}{\max_{(i,j)}\{Z^*\}} \times 100\right) \tag{C.4}$$

The procedure is similar for obtaining daily-frequency data over the entire available period. We first create a rescaled weekly SVI series using the steps outlined above, and then use this rescaled weekly series to construct a correction factor that creates a common index across all of the subperiods of daily observations. Our daily time series of Google search activity is therefore an "index within an index."

Figure C.1 plots the resulting raw, daily SVI series for searches of "public pension payment," excluding searches containing "fees," for our sample time period of April 2011 to October 2014. The dashed red lines indicate scheduled payment dates. Even in the raw data, there is a clear uptick in search activity on payment dates, even when a date is rescheduled from the default 15th of the month. There are also clear spikes in search activity around announcements of reforms to JPS that would affect the determination of benefits. We control for policy announcement effects to isolate the response of searches to the payment schedule in the next subsection.

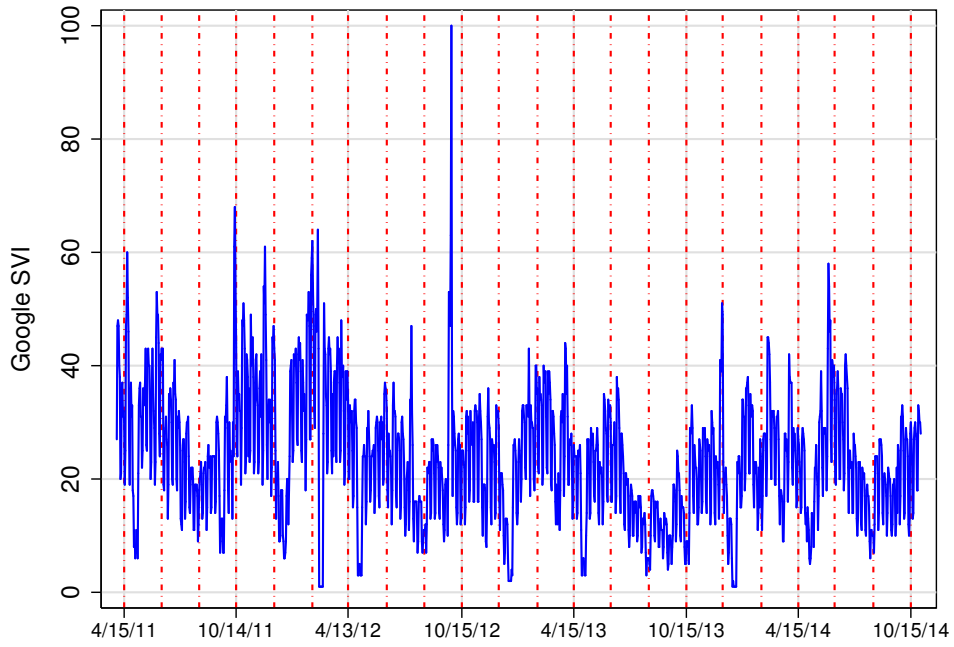
C.2 Estimation Procedure

We run time series regressions of the following form using Google SVI for "public pension payments" relative to average daily SVI as the outcome variable:

$$\widetilde{SVI_t} = \sum_{j=-7}^{+7} \beta_j \cdot Payday_{t+j} + \gamma \cdot t + \delta_{dow} + \phi_{wom} + \psi_{my} + \xi_h + \alpha_p + \epsilon_t \tag{C.5}$$

where we control for the full set of time fixed effects capturing cyclicality in searches within the month (ϕ_{wom}), within the week (δ_{dow}), and within the year (ψ_{my}), as well as holiday effects ξ_h. The linear time trend $\gamma \cdot t$ accounts for the secular increase in internet search activity over the time period. The dummy α_p equals unity on dates when the government first

FIGURE C.1. Raw Daily Google Searches for "Public Pension Payments"



Notes: The figure displays the daily time series of the Japanese Google SVI for "public pension payments." We construct the index according to the steps summarized in equations (C.3) and (C.4). Dashed red lines indicate scheduled pension payment dates during our sample period for the scanner data.

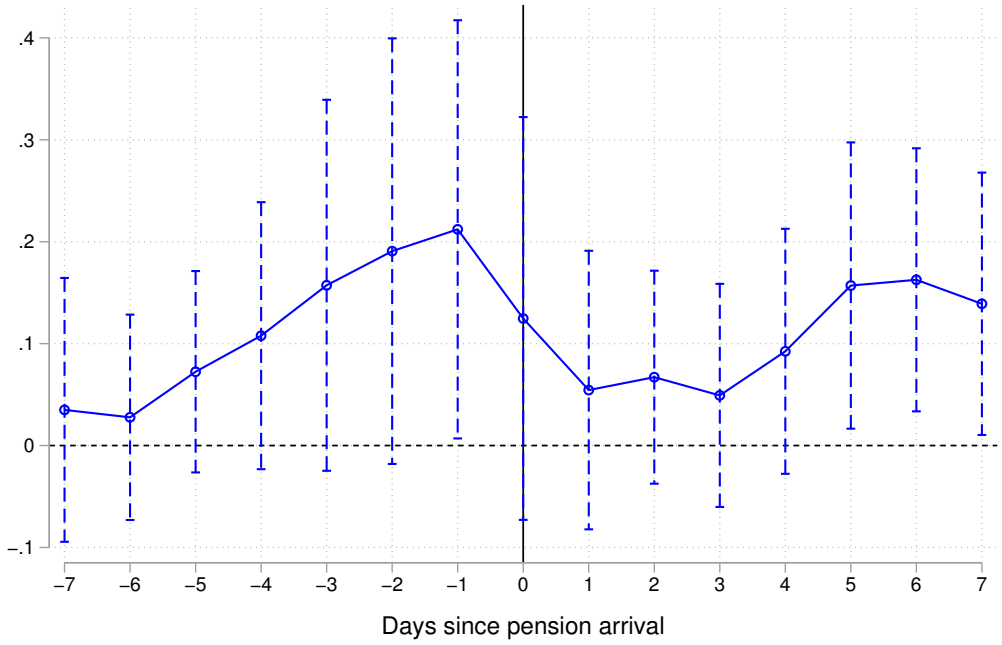
announced policy changes to the pension system, such as changes to the formula determining benefits and the normal retirement age that affected some cohorts; such policy changes are not directly related to the frequency or delivery schedule for payments. However, failing to account for policy announcements could lead to a spurious correlation between payment dates and search activity, as several announcements occur close to scheduled delivery dates.[12]

Figure C.2 plots the estimated coefficients $\widehat{\beta}_j$ which capture search behavior around scheduled pension paydays. Panel A does this for the time series (April 2011 to October 2014) that overlaps with our retail panel, and Panel B extends the time series of Google data (January 2004 to December 2018). In both cases, there is a clear spike in search behavior directly prior to payday, and search activity declines thereafter. This response is equal to 21% above the average daily level of searches on the day before payday in the shorter time series, and 5% above average in the full time series. In the full sample, the decline in searches is sharper, with search volume returning to the average level within two days after a payday. These findings support the notion that individuals are highly attentive to even small deviations of delivery dates from the benchmark 15th of the month.

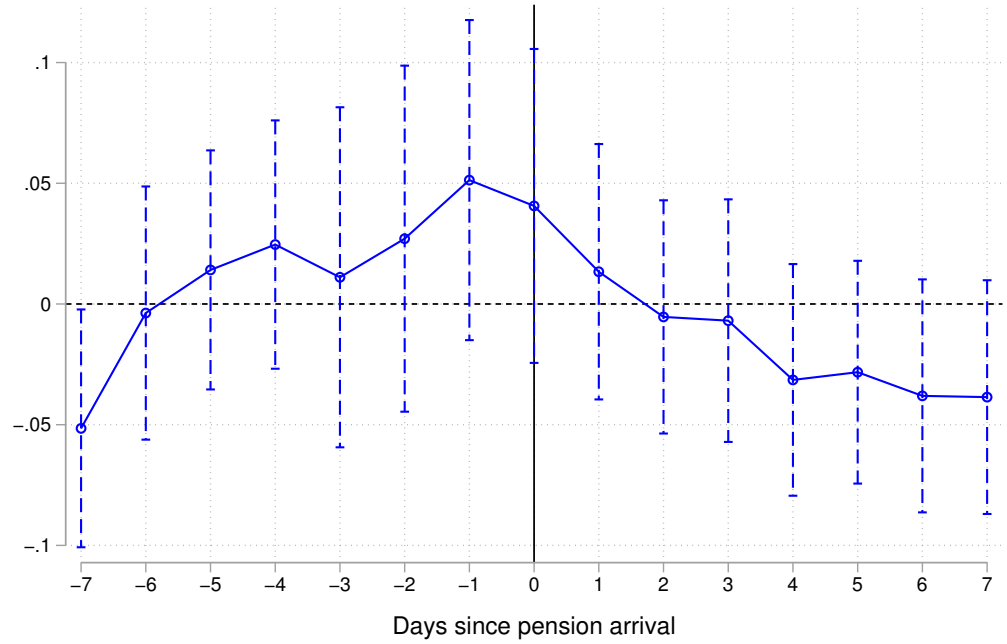
[12]We obtained a list of pension program announcements from the official JPS website: http://www.nenkin.go.jp/oshirase/taisetu/index.html. From this list of 58 announcements during our sample time period, we also tried excluding announcements that were purely clerical in nature (e.g. new application form for dependent beneficiaries), which left us with only three substantive policy change dates. In either case, the estimated coefficient on the dummy α_p was statistically insignificant.

FIGURE C.2. Google Searches for "Pension Payment" around Payday

A. Sample Time Period: April 2011 – October 2014



B. Full Time Period: February 2004 – December 2018



Notes: The figure plots in blue the event study coefficients $\widehat{\beta}_j$ obtained from estimating regression equation (C.5). The y-axis records the log points increase in the Google SVI for "public pension payments" around a scheduled payday. 95% confidence intervals plotted in red.

D Temporary Sales around Payday

In this appendix, we investigate whether retailers respond to increased demand on pension paydays by engaging in price discrimination through reducing the frequency of temporary sales. To the contrary, we find that, if anything, stores are more likely to offer more generous sales on paydays. Moreover, this conclusion holds for several methods we use to identify temporary sales which are common in the literature on sticky prices in high-frequency retail data. These results complement our analysis in Section 5.2, where we use our counterfactual "last price" indices to estimate that approximately 90% of the observed increase in average store-level prices on payday is due to consumers substituting towards an expanded basket of higher quality goods.

In particular, we employ two sets of approaches to identify temporary sales at the store-good level:

(i) The first approach follows Eichenbaum, Jaimovich, & Rebelo (2011) and Kehoe & Midrigan (2008, 2015), who define the most common price during a certain time interval as the regular price. For each product and date, we define the regular price as the three-month centered rolling mode (42 days on either side of the date, or 85 days in total) of the original store-good price series. In what follows we refer to this method as the "rolling mode" method. As noted in Abe & Tonogi 2010, who also use daily scanner data, using a weekly modal price coincides with the definition of bargain sales which are excluded by Statistics Japan when constructing the official Japanese CPI. We obtain qualitatively similar results when we instead use a 7-day rolling mode to characterize temporary sales.

(ii) Our second approach is the sales filter proposed by Nakamura and Steinsson (2008) that classifies V-shaped patterns (i.e. symmetric dips and rebounds) in the price data as temporary sales. This V-filter approach has been shown to pick up actual temporary sales confirmed by field economists, such as in the case of the sales flag in the BLS microdata underlying the CPI (Bils & Klenow 2004). Due to the daily frequency of our data, our parameter choices deviate from the literature which has predominantly used either monthly data (Nakamura & Steinsson 2008) or weekly data like the Kilts-Nielsen retail panel (Chahrour 2011). We set the tuning parameters of the algorithm to be $L = K = J = 42$, or equivalently, we search for V-shaped pricing patterns over a 42-day (1.5 months) window as in Sudo et al. (2018). In what follows we refer to this procedure as the "V-shaped filter."

For each product within a store, we compute the frequency of temporary sales and the average rate of discount. We define the discount rate for a good at date t as $d_t = 1 - p_t/r_t$, where p_t and r_t denote, respectively, the actual price of the good and the regular price computed over the interval containing date t. To be concrete, we compute a store-level

average temporary sales frequency $\overline{f_s}$ and discount rate $\overline{d_s}$ via the following expressions:

$$\overline{f_s} = \frac{1}{|T|} \sum_{t \in T} \left(\frac{1}{|\Omega_s|} \sum_{k \in \Omega_s} \mathbb{1}_t \{ p_{k,s,t} < r_{k,s,t} \} \right) \tag{D.1}$$

$$\overline{d_s} = \frac{1}{|T|} \sum_{t \in T} \left(\frac{1}{|\Omega_s|} \sum_{k \in \Omega_s} \left(1 - p_{k,s,t} / r_{k,s,t} \right) \right) \tag{D.2}$$

where $|T|$ is the number of dates contained in the set T, which may correspond to a set of either paydays or non-paydays in our sample. Ω_s consists of goods k which are sold in store s during our sample time period on both paydays and non-paydays.

In Figure D.1, we plot the kernel density of the store-level average frequency of sales and the average rate of discount on paydays vs. non-paydays. We restrict the sample to goods that are sold both in paydays and non-paydays (about 92% of the original sample). Overall, across all goods we find that stores are equally likely to offer sales on payday vs. non-paydays, but conditional on offering a sale, the discount is likely to be larger on paydays. Under the rolling mode filter, the average payday discount is 0.65 p.p. higher than the average non-payday discount (p-value = 0.000).

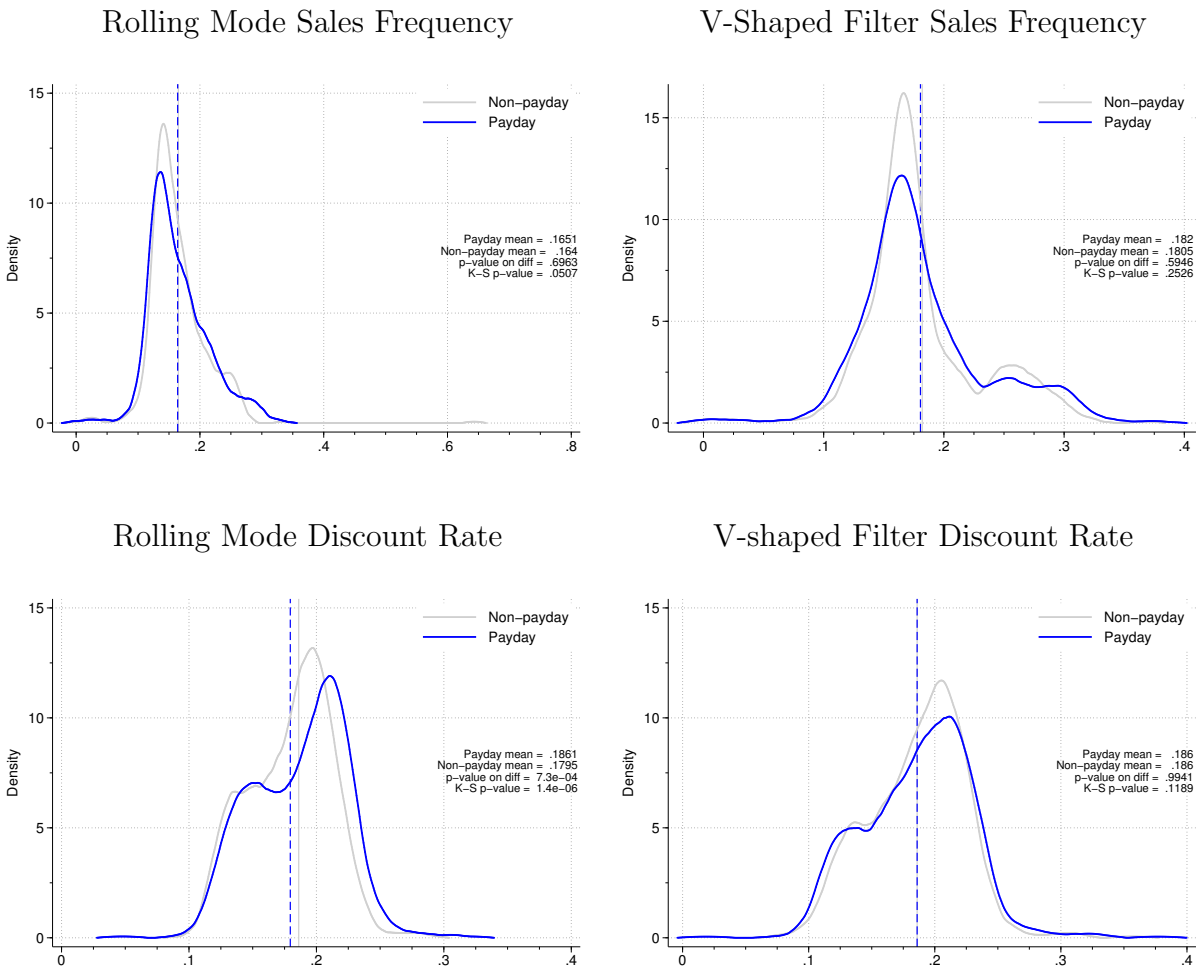
Figure D.2 displays the densities of store-level average sales frequency and discount rates for four-digit goods whose average prices are above the median price within a four-digit goods code, while Figure D.3 does the same for below-median price goods.[13] Dividing goods into above and below-median prices within a four-digit category is analogous to comparing, for example, the propensity of stores to offer discounts on a blended whiskey (below-median price) versus a single-malt whiskey (above-median price) produced by the same brand. As shown in Bils & Klenow (2001), prices within a goods category are increasing in household expenditures – that is, quality Engel curves slope upward. Splitting fine categories of goods into price quantiles is therefore analogous to sorting goods according to quality, or the income profile of their consumers.

We uncover stark differences in retailer pricing strategies by product quality that get partially washed out when we look at temporary sales pooled across all goods. For above-median price goods and across both sales filters, average temporary sales frequency is about 1 p.p. lower on paydays compared to non-paydays, while discounts are between 0.5 p.p. (rolling mode) to 1 p.p. (V-shape) higher on non-paydays. In other words, Figure D.2 shows that for high-quality goods, stores offer fewer sales promotions on paydays, and the ones they do offer feature less generous discounts. This pattern is completely reversed in Figure D.3 where we examine sales discounts for below-median price goods. Average temporary sales frequency is about 1.5 p.p. higher on paydays, while discounts are between 0.8 p.p. (V-shape) and 1.5 p.p. (rolling mode) higher on paydays.

That stores offer more aggressive sales promotions for lower-quality products on paydays is consistent with the idea that price discriminating retailers can generate profit by lowering

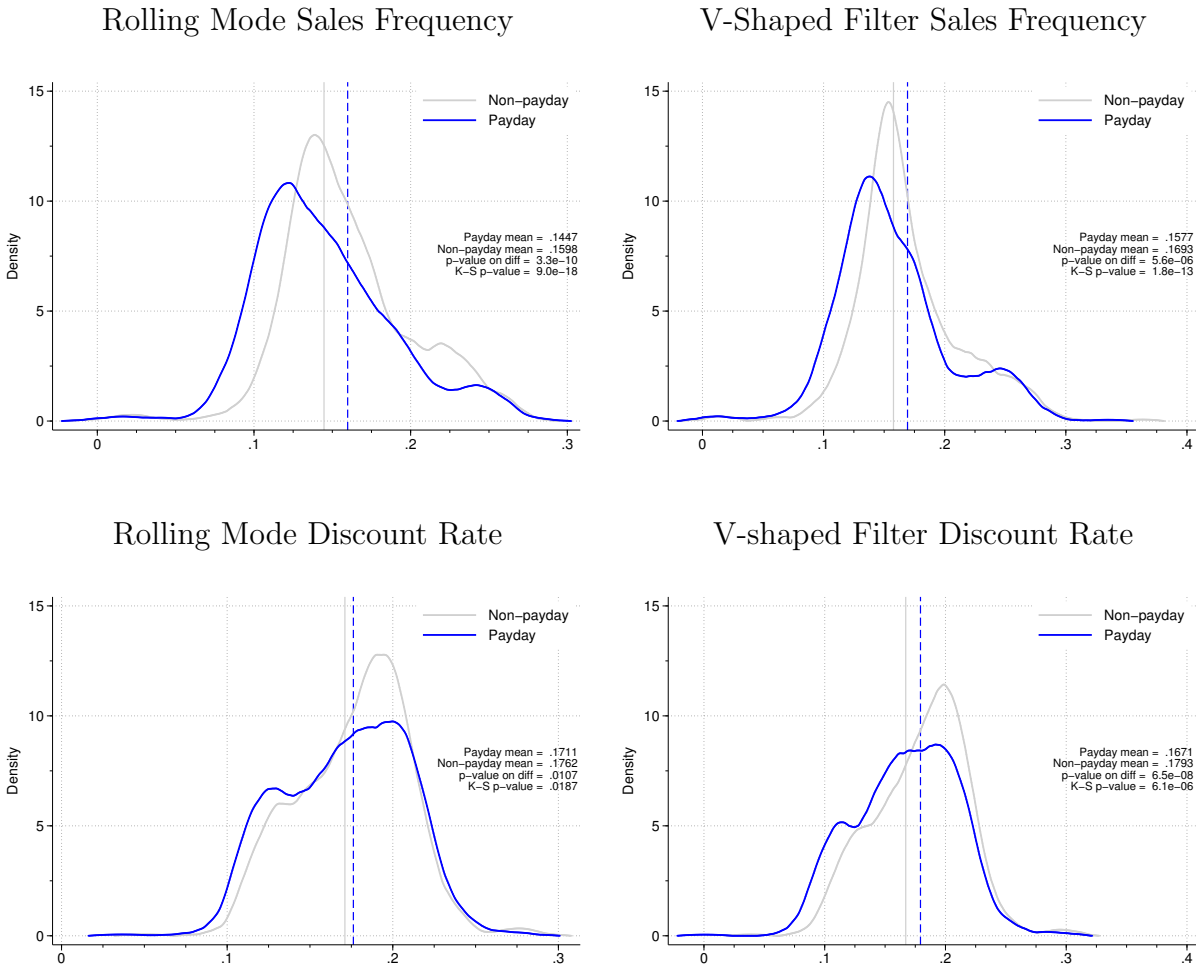
[13] See Appendix B for details on how goods categories are defined in the scanner data. For example, four-digit categories within the two-digit category of "raw meat and poultry" include "wagyu" and "eggs."

FIGURE D.1. Store-Level Temporary Sales Frequency and Discounts



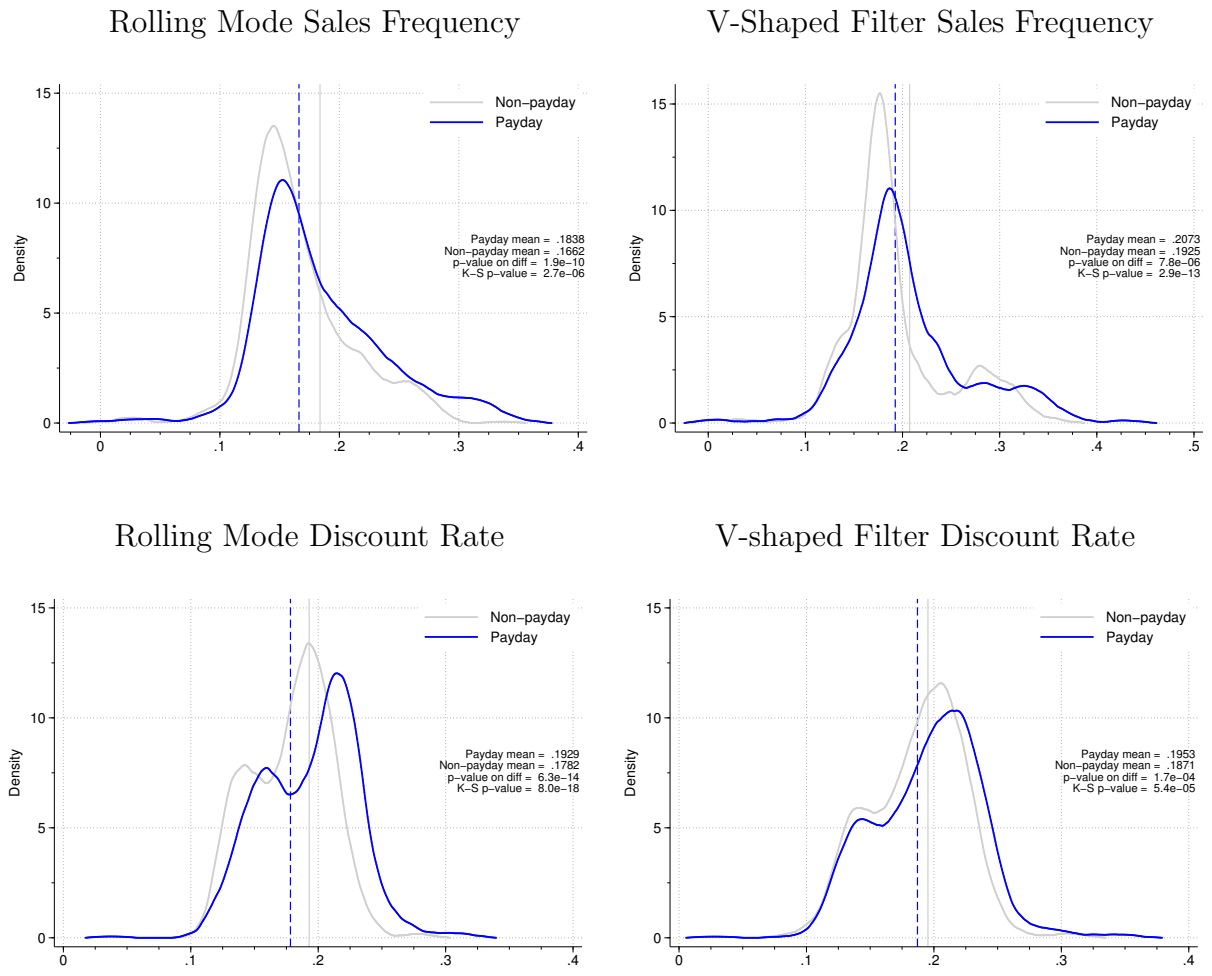
Notes: The top two panels display the kernel densities for average store-level temporary sales frequencies pooling all products sold at a store across both paydays and non-paydays. A frequency of 0.2, for example, implies the average product at a store goes on sale once every $1/0.2 = 5$ days. The bottom two panels perform the same exercise for the average store-level discount rate, where the discount rate is 1 minus the ratio of the current price to the regular price. The left-hand side panels show the frequencies and discount rates under the rolling mode filter, while the right-hand side panels show the distributions when we use the V-shaped filter to identify sales. In both algorithms we search for temporary sales over a 42-day window on either side of a calendar date. Solid grey vertical lines indicate the mean daily frequency or discount rate across stores on non-paydays, while blue dashed lines show the mean across stores on paydays. The K-S p-value shows the two-sided exact p-value from a Kolmogorov-Smirnov test of equality for the payday vs. non-payday distributions.

FIGURE D.2. Temporary Sales Frequency and Discounts on Above-Median Price Goods



Notes: The top two panels display the kernel densities for average store-level temporary sales frequencies, including only products which have an above-median average price within their four-digit goods category. A frequency of 0.2, for example, implies the average product at a store goes on sale once every $1/0.2 = 5$ days. The bottom two panels perform the same exercise for the average store-level discount rate, where the discount rate is 1 minus the ratio of the current price to the regular price. The left-hand side panels show the frequencies and discount rates under the rolling mode filter, while the right-hand side panels show the distributions when we use the V-shaped filter to identify sales. In both algorithms we search for temporary sales over a 42-day window on either side of a calendar date. Solid grey vertical lines indicate the mean daily frequency or discount rate across stores on non-paydays, while blue dashed lines show the mean across stores on paydays. The K-S p-value shows the two-sided exact p-value from a Kolmogorov-Smirnov test of equality for the payday vs. non-payday distributions.

FIGURE D.3. Temporary Sales Frequency and Discounts on Below-Median Price Goods



Notes: The top two panels display the kernel densities for average store-level temporary sales frequencies, including only products which have a below-median average price within their four-digit goods category. A frequency of 0.2, for example, implies the average product at a store goes on sale once every $1/0.2 = 5$ days. The bottom two panels perform the same exercise for the average store-level discount rate, where the discount rate is 1 minus the ratio of the current price to the regular price. The left-hand side panels show the frequencies and discount rates under the rolling mode filter, while the right-hand side panels show the distributions when we use the V-shaped filter to identify sales. In both algorithms we search for temporary sales over a 42-day window on either side of a calendar date. Solid grey vertical lines indicate the mean daily frequency or discount rate across stores on non-paydays, while blue dashed lines show the mean across stores on paydays. The K-S p-value shows the two-sided exact p-value from a Kolmogorov-Smirnov test of equality for the payday vs. non-payday distributions.

prices during periods of peak demand to attract more price-sensitive "bargain hunter" types without affecting their profits from "brand loyalists" who always buy the higher quality version of the same good (Chevalier & Kashyap 2019). In our main analysis, we show that over 90% of the observed increase in store-level average prices is due to consumer substitution towards discretionary purchases of higher-end goods, or "splurge" goods. Given that retailers keep prices on splurge goods high during paydays by forgoing temporary sales, it is unlikely that the government's choice of payment frequency would have an additional effect on intra-cycle consumption patterns through a supply-side channel. However, in Appendix A we consider some extensions of our model with monopolistic firms offering a price menu which depends on the government's payment interval choice.

E Japanese Pension System Branch Offices

This appendix offers details on the data underlying our estimation in Section 6.1 of the marginal cost to administering the Japanese Pension System (JPS). We downloaded the list of JPS branch office locations and overlaid modern municipal boundaries, resulting in the map pictured in Figure E.1.[14] Out of the 663 municipalities reporting public expenditures dating back to the 1980s, roughly 1/3 (N = 239) contain a JPS branch office. Branch locations do not appear to be any more geographically isolated from non-branch office locations. An exception is the Tokyo and Osaka metropolitan areas, where each city ward contains its own JPS branch. Once we exclude major cities from the sample (column 6 of Table 3), we do not find any significant uptick in costs, suggesting that our baseline calibrations of the optimal frequency lead to an upward bias in the recommended payment interval lengths.

Table E.1 shows that, as of the quincennial Census year (1985) preceding the 1988 reform which increased the frequency of pension payments from every three to two months, branch office cities tend to be more population dense, have a younger population, and have achieved higher per capita incomes. However, these differences are economically small, and the two sets of localities are balanced on observables which are key to the administration of the pension system – namely, the ratio of government expenditures to revenues, and welfare spending on the local elderly population. The result is that in Figure 8, prior to the 1988 reform, the time series for spending per elderly resident are not only parallel, but lie directly on top of each other. Our marginal cost estimates imply small costs per claimant (6,413 to 13,770 JPY) of shifting towards a bimonthly payment system even after we include population or per capital income $\times$ time fixed effects (columns 3 and 5 of Table 3).

We measure administrative costs per claimant, $\mu_{j,t}$ in equation (6.1), as the costs per person over 65 associated with providing elderly welfare benefits. We obtain spending and revenue items for each modern municipality from the Cabinet Office's local public finance database.[15] While local JPS branch offices are not directly responsible for remitting payments

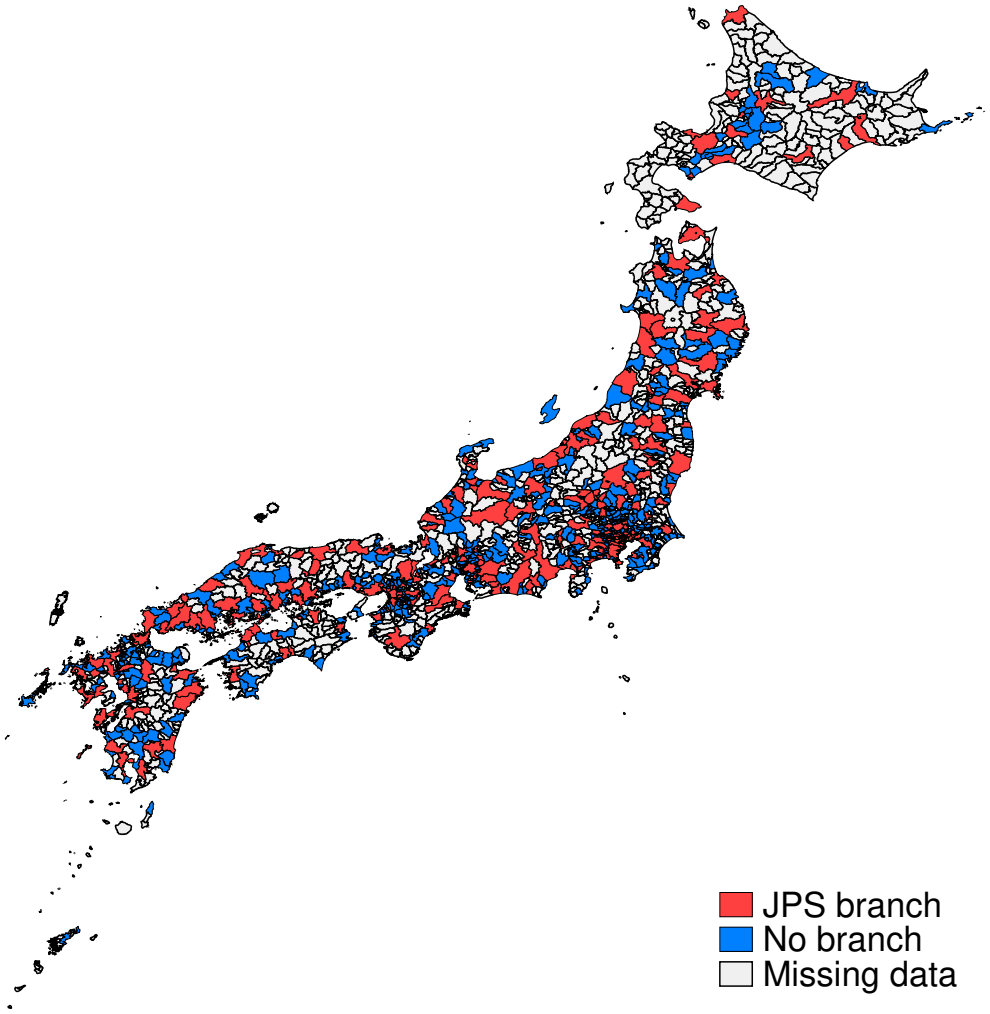
[14]The list of branch office locations is maintained here: https://www.nenkin.go.jp/international/aboutjps/offices.html.

[15]The balance sheet line items can be downloaded at the MIERUKA database, accessible here: https://wwwb.cao.go.jp/ittaikaikaku/mieruka/.

to recipients, they do incur administrative costs related to processing applications, computing tax withholding from payments, and arbitrating disputes and fraud cases.[16] Unfortunately, the municipal balance sheets do not separate spending on local elderly welfare benefits from costs incurred by supporting the JPS branch office system.

[16]See, for instance, official descriptions in the 2022 version of the "Japanese Pension Service and its Operation," https://www.nenkin.go.jp/international/aboutjps/operation.files/about_jps_operation.pdf.

FIGURE E.1. Map of Japanese Pension System Branch Office Locations



Notes: The map displays municipalities with (red) and without (blue) Japanese Pension System (JPS) offices as of 1980, when our local government balance sheet data are first widely available. We exclude from our city-level panel any local jurisdictions with gaps in public expenditure data around the 1988 reform which reduced the payment interval length from three months to two months; such cities are indicated on the map in gray ("missing data"). In our analysis, we impose modern municipal boundaries defined according to the city code crosswalk available through RIETI (Kondo 2019). Our estimates of the slope of the administrative cost function for the JPS are nearly unchanged when we instead impose historical 1980 municipal boundaries to assign treatment status.

Table E.1. Summary Statistics for JPS Branch vs. Non-Branch Office Cities

	Branch (N = 239)	Non-branch (N = 424)	Difference	p-value
Log Census population	12.16	11.07	1.09	0.00
	(0.06)	(0.03)	(0.06)	
CBD population density (1000s/km^2)	7.66	5.11	2.55	0.00
	(0.46)	(0.15)	(0.38)	
Fraction population > 65 y.o. (%)	10.71	11.17	−0.46	0.09
	(0.17)	(0.18)	(0.26)	
Fraction population > 75 y.o. (%)	4.05	4.26	−0.21	0.07
	(0.07)	(0.08)	(0.12)	
$\%\Delta^{75-85}$ population > 65 y.o.	43.00	46.73	−3.73	0.02
	(0.94)	(1.12)	(1.65)	
Fraction female residents (%)	51.14	51.09	0.05	0.63
	(0.00)	(0.00)	(0.00)	
Fertility rate	2.33	2.28	0.05	0.04
	(0.02)	(0.02)	(0.03)	
Log per capita income	7.82	7.79	0.03	0.01
	(0.01)	(0.01)	(0.01)	
Labor force participation rate (%)	50.24	49.71	0.53	0.06
	(0.21)	(0.16)	(0.28)	
Unemployment rate (%)	3.48	3.08	0.40	0.00
	(0.09)	(0.06)	(0.10)	
Ratio of govt. expenditures to revenues	0.97	0.97	0.00	0.89
	(0.00)	(0.00)	(0.00)	
Log welfare spending per person > 65 y.o.	4.10	4.05	0.05	0.18
	(0.03)	(0.02)	(0.04)	
Log welfare spending per person > 75 y.o.	5.08	5.03	0.05	0.19
	(0.03)	(0.02)	(0.04)	

Notes: The table provides the mean and standard deviation, the unconditional difference in means and standard error, and p-value on the two-sided t-test of the difference between municipalities with a branch office of the Japanese Pension System (JPS) and those without a branch office. All non-growth rate variables are measured as of the pre-reform period in 1985. Population variables come from the quinquennial Census. The fertility rate is defined here as the number of live births per 100 females of child-bearing age. To obtain per capita income (in 1,000s of JPY), we use the Cabinet Office local statistics for taxable income and divide by total 1985 Census population. Government expenditure ratios and unemployment rates also come from the Cabinet Office local statistics. To compute these statistics, we impose modern municipal boundaries using the historical city code crosswalk available through RIETI (Kondo 2019).

F Additional Results on Payday Responses

In this appendix, we provide supplemental results on consumer and retailer responses to payday in our empirical application to the Japanese Pension System. We document four new results:

1. Spending responses follow similar trajectories regardless of shoppers' store visit frequency, with smaller payday spending spikes for most goods subcategories.
2. The spike in overall expenditures on payday is due to substitution towards more expensive varieties of goods in splurge categories.
3. Retailers largely fail to price discriminate by goods subcategory, with the exception of prepared foods for which the response accounts for only 20% of the overall spike in store-level prices on prepared foods.
4. Pension recipients spend *more* on perishable goods (i.e. instantaneous consumption) after crossing the early (age 60) and normal (age 65) retirement age thresholds. The third result casts doubt on the idea that spending responses are driven by elderly consumers having insufficient savings as they enter retirement.

F.1 Expenditure Responses by Category, Trip Frequency, & Income Decile

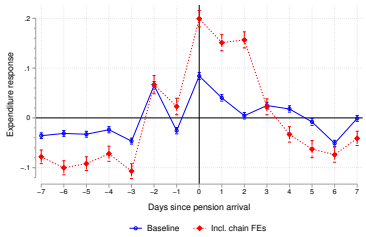
Figure F.1 reproduces our main event study Figure 3, but with shopper spending broken down by goods subcategory. Spending around payday is concentrated in splurge goods: prepared foods, sweets/desserts, alcohol, fresh fish and meat, and other processed foods which primarily contain snack items. Conditional on chain fixed effects which account for retailer-specific sales promotions and inventory shocks, spending responses are stronger on the scheduled payment date, and for most categories, spending reverts to the mean after three days since payment arrival.

Figure F.2 re-estimates the subcategory spending event studies using a sample of very frequent shoppers for whom average monthly grocery spending matches statistics reported in the Family Income and Expenditure Survey (FIES) over the same time period, 2011-2014. See Appendix B for detailed summary statistics and information on how we created this shopper subsample. While the magnitude of the payday spending spike is smaller across most subcategories, the trends around payday closely mimic those exhibited in Figure F.2 when we use our less restrictive sample of weekly shoppers. As the specifications with store chain and shopper-specific day-of-week fixed effects show, this effect is not driven by differences in store brand preferences between more vs. less frequent shoppers.

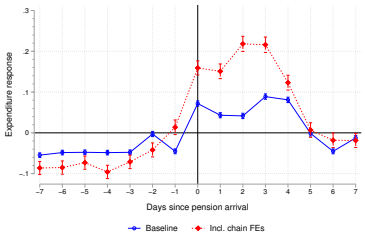
Complementing our results in Section 4.4, Table F.1 reports perishable ("raw foods") payday expenditures by permanent income decile, where we proxy for permanent income using average total pay cycle expenditures within each regular shopper's history (Kueng 2018). We uncover a similar pattern regardless of whether we focus on the intensive (i.e. how

FIGURE F.1. Weekly Shoppers' Payday Subcategory Spending Responses

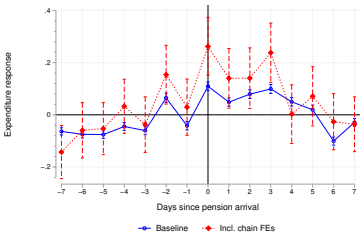
(a) Prepared Foods



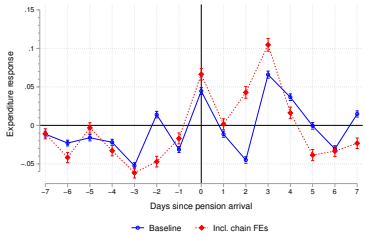
(b) Sweets/Desserts



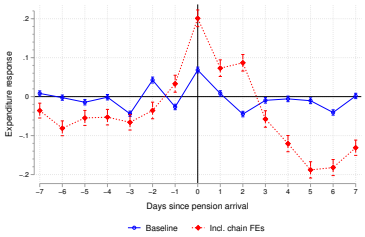
(c) Alcohol



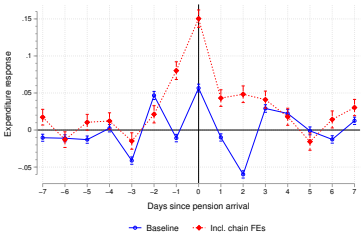
(c) Fresh Produce



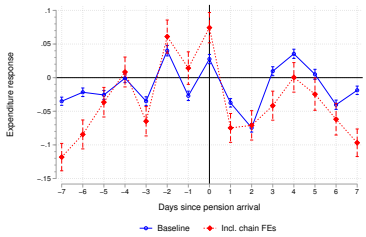
(d) Fresh Fish



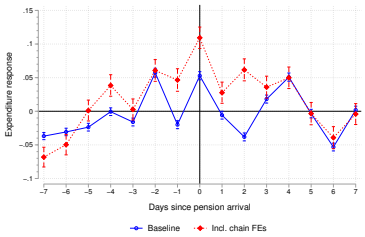
(e) Meat & Poultry



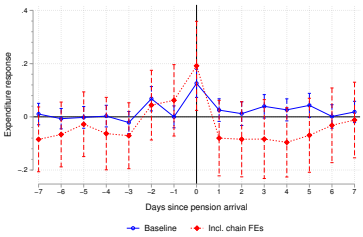
(f) Grains



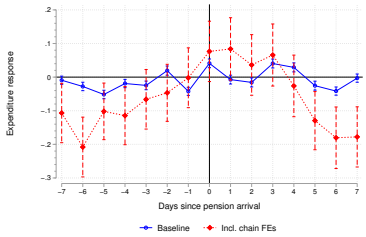
(g) Non-alcoholic beverages



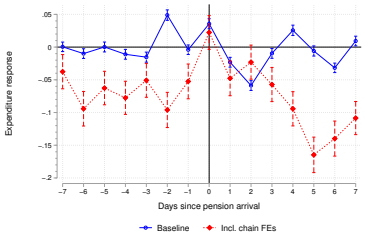
(h) Tobacco



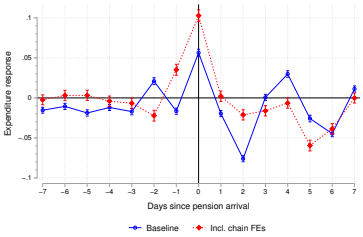
(i) Processed Fruits/Vegetables



(j) Preserved Fish

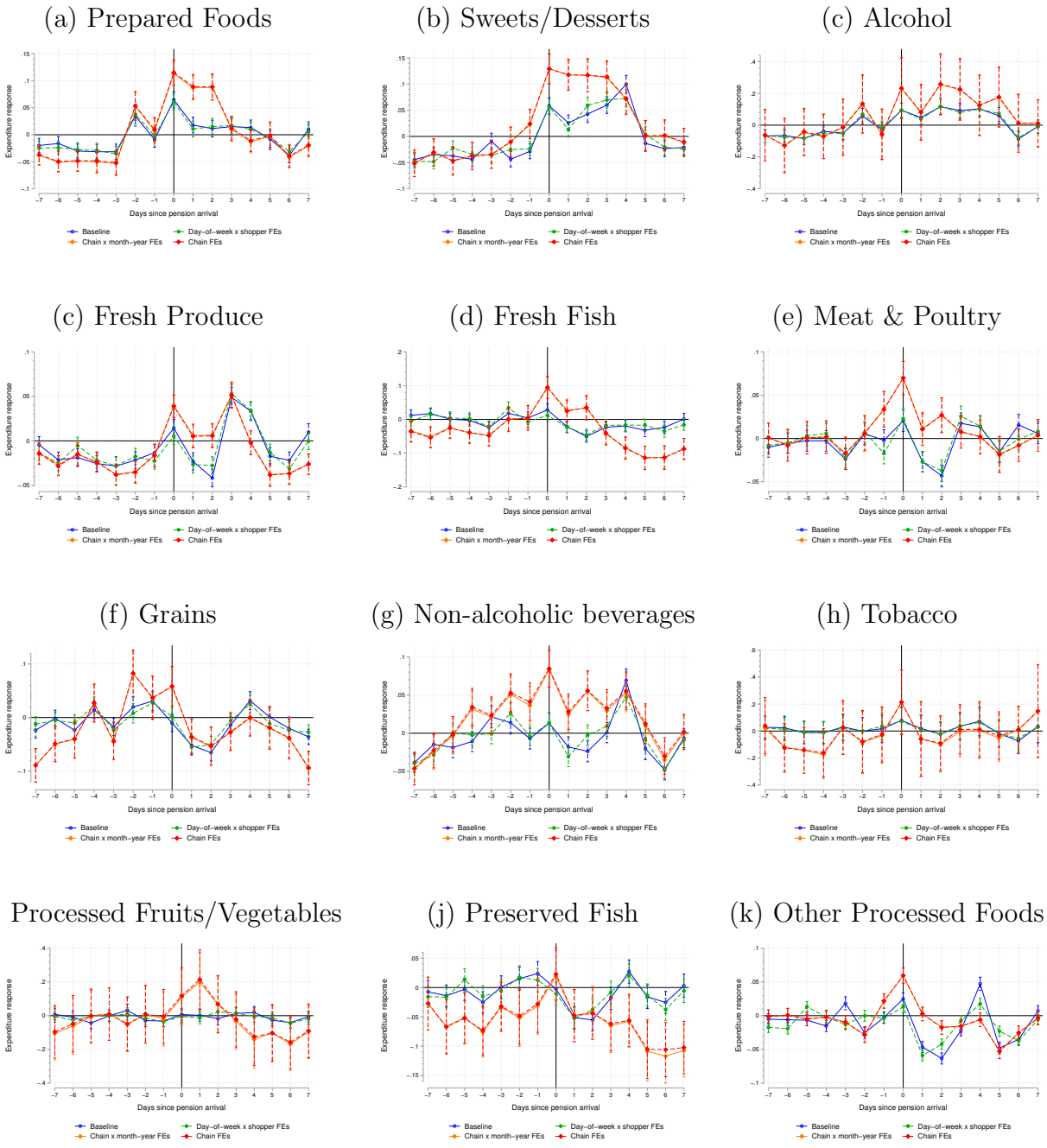


(k) Other Processed Foods



Notes: Each panel plots event study coefficients $\widehat{\beta}_j$ obtained from estimating regression equation (4.1) for a major goods expenditure category over a panel of weekly shoppers which remains fixed across categories. Point estimates in red obtained from a version of equation (4.1) which includes store chain fixed effects to account for variation in spending patterns due to retailer-specific pricing decisions. The y-axis records the percent increase in expenditures relative to average daily expenditures within that category over the panel. We winsorize the top 1% of total daily expenditures within each subcategory. Bars indicate 99% confidence intervals, with standard errors obtained from clustering by shopper ID. See Appendix B for details on how we sort goods into each expenditure category.

FIGURE F.2. Very Frequent Shoppers' Payday Subcategory Spending Responses



Notes: Each panel plots the event study coefficients $\widehat{\beta}_j$ obtained from estimating regression equation (4.1) for a major goods expenditure category over a panel of very frequent shoppers which remains fixed across categories. We construct the panel of very frequent shoppers using the procedures outlined in Appendix B. Point estimates in green interact the shopper fixed effects η_i with the day-of-week fixed effects δ_{dow} to account for individuals' preferences for shopping on particular days of the week. Point estimates in orange obtained from a version of equation (4.1) which adds store chain $\times$ month-year fixed effects to account for variation in spending patterns due to time-varying retailer-specific pricing decisions. The specification in red is estimated off the intensive margin sample of shopper trips. The y-axis records the percent increase in expenditures relative to average daily expenditures within that category over the panel. We winsorize the top 1% of total daily expenditures within each subcategory. Bars indicate 99% confidence intervals, with standard errors obtained from clustering by shopper ID.

Table F.1. Payday Expenditure Responses by Permanent Income Decile

	1st	2nd	3rd	4th	5th	6th	7th	8th	9th	10th
Intensive margin	0.106***	0.097***	0.088***	0.074***	0.069***	0.054***	0.052***	0.048***	0.035***	0.031***
	(0.011)	(0.009)	(0.008)	(0.007)	(0.006)	(0.006)	(0.005)	(0.004)	(0.003)	(0.003)
Intensive margin w/controls	0.0104***	0.094***	0.085***	0.073***	0.067***	0.051***	0.050***	0.046***	0.033***	0.029***
	(0.011)	(0.009)	(0.008)	(0.007)	(0.006)	(0.006)	(0.005)	(0.004)	(0.003)	(0.003)
Total response	0.069***	0.062***	0.047***	0.045***	0.036***	0.035***	0.026***	0.029***	0.022***	0.003
	(0.005)	(0.005)	(0.005)	(0.005)	(0.004)	(0.004)	(0.004)	(0.004)	(0.003)	(0.003)
Extensive margin	0.008***	0.007***	0.003*	0.003**	0.002	0.001	−0.002+	−0.002+	−0.002+	−0.012***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)

Notes: The table reports expenditure responses of regular shoppers of age 65 or older to payday, sorted by decile of average total pay cycle expenditures. Each column number refers to results for a decile of average total expenditures over bimonthly pay cycles. Each row refers to a version of our baseline specification in equation (4.1) with a different measure of expenditures as the outcome variable. Intensive margin regressions in row 1 condition on the individual visiting the store on a given date. Regressions in row 2 record intensive margin responses but control for the number of days in between shopping trips. For total response regressions, the dependent variable is expenditures on raw foods relative to average daily expenditures on raw foods. Extensive margin refers to regressions where the outcome variable is a dummy equal to one if the shopper visits a store in our sample on a given date. Hence, coefficients reported for the extensive margin regressions represent changes in probability that a shopper visits a store as a result of payday. We winsorize the top 1% of daily expenditures. Robust standard errors clustered by shopper ID in parentheses. $^{***}p < 0.001, ^{**}p < 0.01, ^{*}p < 0.05, +p < 0.1$

much the shopper spends conditional on visiting a store), extensive margin (i.e. whether a shopper visits a store), or the total spending response (intensive × extensive margin). Spending responses are weakly, negatively monotone in permanent income – with a much flatter gradient for shoppers with above-median permanent income. The result also holds even if we control for the number of days in between shopping trips (row 2 of the table).

F.2 Accounting Decomposition of Store-Level Price Changes

We show how movements in the geometric average price index in equation (5.1) can be decomposed into three components: (i) price changes enacted by retailers for a basket of commonly transacted goods, (ii) consumers expanding the set of goods they purchase within the common basket, or a "variety" response, and (iii) consumers substituting towards goods of different prices outside the common basket. Our analysis in Section 5 separates (i) from the sum of (ii) and (iii) to document negligible retailer price discrimination in response to pension payment receipt.

In Appendix D, we further analyze how effect (i) can originate from changes in the incidence of temporary sales, or changes in the discount rate in temporary sales. To the extent that there is a positive effect due to channel (i), we find this is driven by slightly less frequent and less generous discounts for above-median price barcodes within a 4-digit category. Sales frequencies and discount rates follow the opposite pattern for below-median price barcodes.

The average price index for store s on date t is given by:

$$\Phi_{s,t} = \frac{1}{n_{s,t}} \sum_{k \in \Omega_{s,t}} \log p_{k,s,t} \tag{F.1}$$

where $\Omega_{s,t}$ is the set of barcodes k transacted in that store on a given date. Denote the cardinality of this set $n_{s,t} = |\Omega_{s,t}|$. The store-level inflation rate is then:

$$\Delta\Phi_{s,t} = \frac{1}{n_{s,t}} \sum_{k \in \Omega_{s,t}} \log p_{k,s,t} - \frac{1}{n_{s,t-1}} \sum_{k \in \Omega_{s,t-1}} \log p_{k,s,t-1} \tag{F.2}$$

Partition the set of transacted barcodes into three subsets:

$$\begin{aligned}
\Omega^{*}_{s,t,t-1} &= \Omega_{s,t} \cap \Omega_{s,t-1} \\
\Omega^{new}_{s,t,t-1} &= \Omega_{s,t} \backslash \Omega^{*}_{s,t,t-1} \\
\Omega^{old}_{s,t,t-1} &= \Omega_{s,t-1} \backslash \Omega^{*}_{s,t,t-1}
\end{aligned}$$

Hence, Ω^{*} is the set of commonly transacted goods between dates $t-1$ and t, Ω^{new} is the set of goods only purchased in t, and Ω^{old} the set of goods only purchased in $t-1$.

Using this set partition, we can then decompose store-level price inflation into the three

aforementioned effects:

$$\Delta\Phi_{s,t} = \left(\frac{1}{n_{s,t}}\sum_{k\in\Omega^*}\log p_{k,s,t} - \frac{1}{n_{s,t-1}}\sum_{k\in\Omega^*}\log p_{k,s,t-1}\right) + \left(\frac{1}{n_{s,t}}\sum_{k\in\Omega^{new}}\log p_{k,s,t} - \frac{1}{n_{s,t-1}}\sum_{k\in\Omega^{old}}\log p_{k,s,t-1}\right) \tag{F.3}$$

$$\implies \Delta\Phi_{s,t} = \underbrace{\frac{1}{n_{s,t}}\sum_{k\in\Omega^*}\Delta\log p_{k,s,t}}_{\text{retail price change of common goods}} + \underbrace{\left(\frac{1}{n_{s,t}} - \frac{1}{n_{s,t-1}}\right)\sum_{k\in\Omega^*}\log p_{k,s,t-1}}_{\text{consumer variety response}}$$
$$+ \underbrace{\left(\frac{1}{n_{s,t}}\sum_{k\in\Omega^{new}}\log p_{k,s,t} - \frac{1}{n_{s,t-1}}\sum_{k\in\Omega^{old}}\log p_{k,s,t-1}\right)}_{\text{consumer substitution towards new goods}} \tag{F.4}$$

Equation (F.3) tells us that store-level inflation is due to the change in the average price index level within Ω^*, plus consumer substitution between old and new goods. Equation (F.4) further decomposes the former term into price setting by the retailer and the change in the number of goods purchased within Ω^* (i.e. variety). The strength of the variety effect is proportional to $\Delta\log n_{s,t}$, which we estimate in Panel B of Figure 6.

Empirically implementing this accounting identity requires us to define the time period $[t-1, t]$ over which we partition the purchase set Ω. Defining $\Omega^*_{s,t}$ as the intersection of daily sets results in very sparse baskets of common goods for some stores and dates with limited traffic. To mitigate this issue, in Section 5 we introduced the concept of a "last price" index so that the time period $[t-1, t]$ used to partition $\Omega_{s,t}$ is distinct from the daily frequency at which we estimate our event study equations.

F.3 Retailer Pricing Responses

Figures F.3, F.4, and F.5 reproduce the event studies in Figure 6 and Figure 7 by goods subcategory. Figure F.3 shows that store-level prices, measured using the geometric average price index in equation (5.1), increased for all goods categories on payday. For the splurge categories of goods (prepared foods, sweets/desserts, alcohol), observed prices remained higher for the entire week after payday. Figure F.4 demonstrates that these pricing patterns mirror consumer substitution towards an expanded set of varieties within each category.

In Figure F.5, we isolate the component of store-level prices due to retailers changing posted prices. We do this using the same procedure described in Section 5.2. That is, we compare the event study coefficients for store-level prices within each category displayed in Figure F.3 and then estimate the same event study equation (5.2) but using a counterfactual "last price" index which substitutes the actual contemporaneous price with the last observed price for each barcode in the odd month preceding the payday. With the exception of prepared foods, the fraction of the hike in store-level prices around payday explained by changes in

retailer behavior is quantitatively small ($< 10\%$), and for non-splurge goods, these pricing responses are close to zero.

Table F.2 separately tabulates for each goods category and for overall spending the fraction of observed store-level pricing responses due to retailer behavior averaged over a $\pm$ 3-day event window, on payday itself, and for two definitions of the last price index: the last-month and last-week indices we previously used in Figure 7. Computing the pricing responses over a 3-day symmetric window around payday allows for the possibility that consumers may receive payments slightly earlier or later than the scheduled date due to differences across banks in the timing of processing deposits. More concretely, we run the two event studies:

$$\Phi_{s,t} = \sum_{j=-7}^{+7} \gamma_{1,j} \cdot Payday_{t+j} + \delta_{dow} + \phi_{wom} + \xi_h + \eta_s + \varphi_{c,my} + \epsilon_{s,t} \tag{F.5}$$

$$\Phi_{s,t}^{last} = \sum_{j=-7}^{+7} \gamma_{2,j} \cdot Payday_{t+j} + \delta_{dow} + \phi_{wom} + \xi_h + \eta_s + \varphi_{c,my} + \upsilon_{s,t} \tag{F.6}$$

where $\Phi_{s,t}$ is given by (F.1), and $\Phi_{s,t}^{last}$ is the counterfactual last price index which takes a geometric average of barcode prices $p_{k,s,last}$ at some time $last$ preceding each payday event. We include retail chain by month-year fixed effects $\varphi_{c,my}$ to account for seasonal sales promotions and supply shocks differentially impacting store inventories. Hence, $\Phi_{s,t}^{last}$ varies over time due to shifts in the composition of the basket of goods purchased, as captured by $n_{s,t}$ in the decomposition (F.4).

In estimating (F.5) and (F.6), we normalize both indices by their average value over the sample period. We then report in Table F.2 the fraction of store-level inflation driven by retailer responses averaged over the week containing the payday:

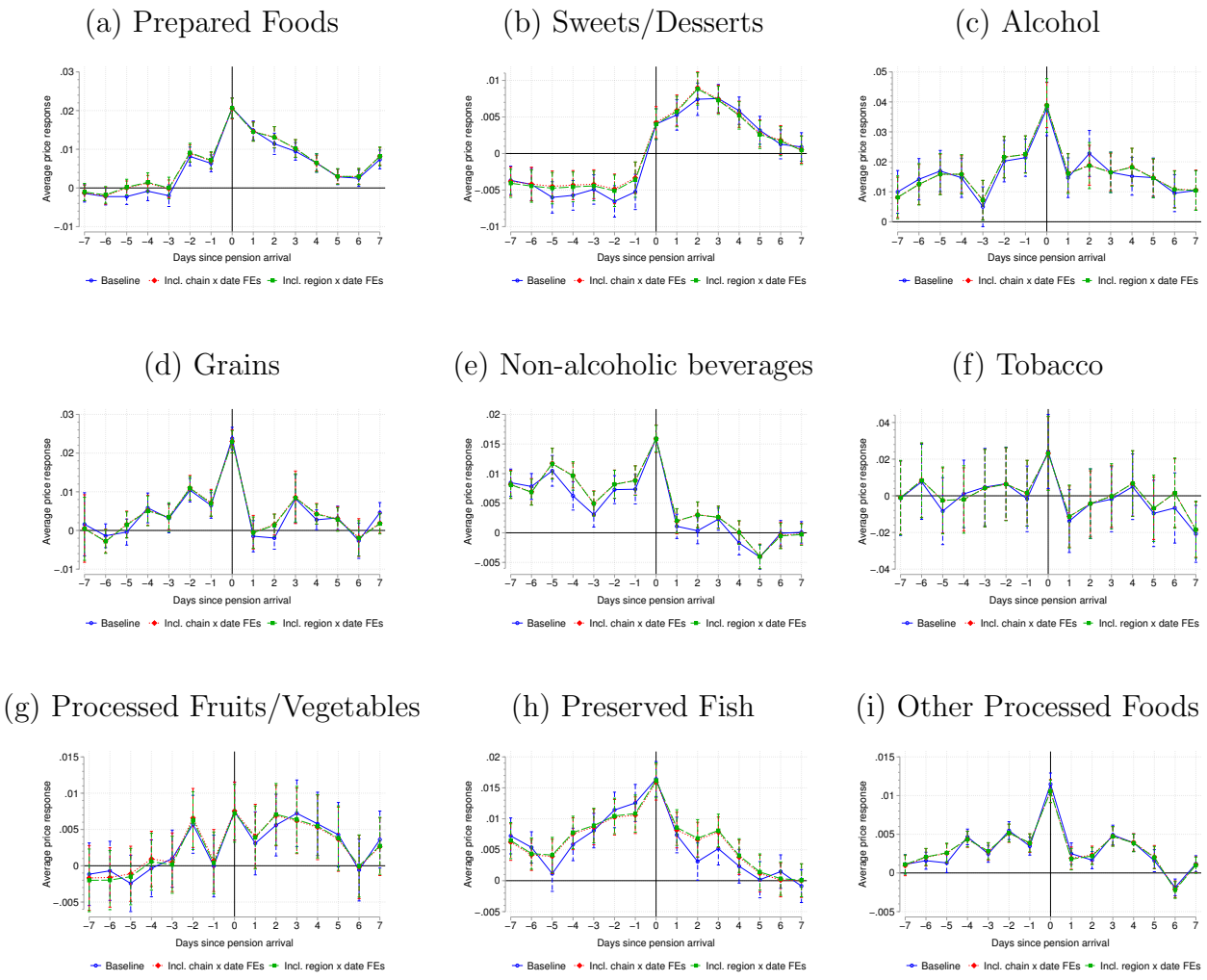
$$\bar{\eta} \equiv \frac{1}{7} \left\{ \sum_{j=-3}^{+3} \frac{\widehat{\gamma}_{1,j} - \widehat{\gamma}_{2,j}}{\widehat{\gamma}_{1,j}} \right\} \tag{F.7}$$

and, analogously, we define $\eta_0 \equiv (\widehat{\gamma}_{1,0} - \widehat{\gamma}_{2,0})/\widehat{\gamma}_{1,0}$ as the payday response. We block bootstrap standard errors at the store level by repeatedly re-estimating expressions (F.5–F.7) over draws of subsamples of stores with replacement.

F.4 Spending Responses around Retirement Age

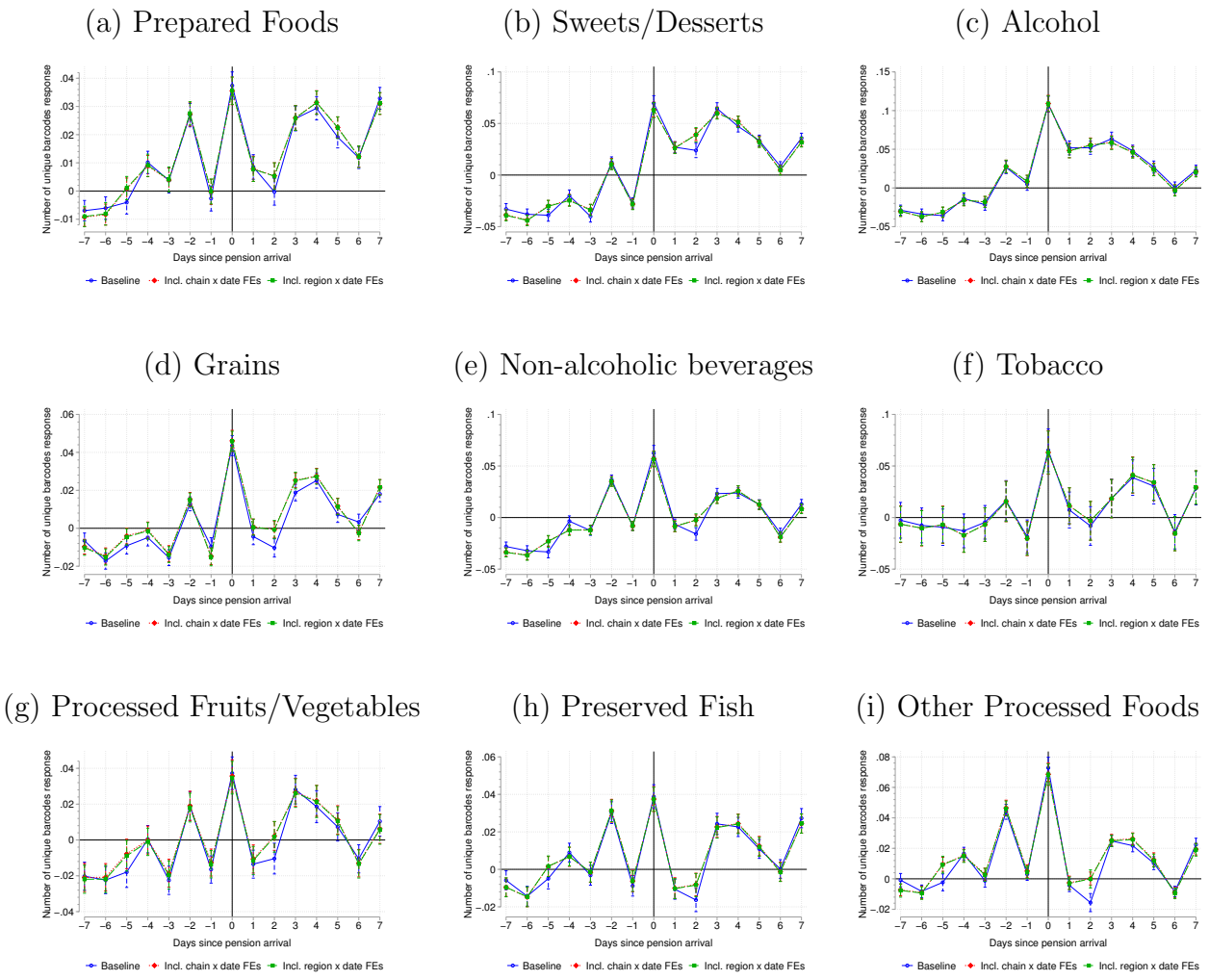
How do spending responses to pension payday vary around the early (60 years old) and normal (65 years old) retirement age thresholds? Answering this question will help clarify two items: (i) the potential role of liquidity constraints for generating the observed spike in overall spending among elderly shoppers on payday, and (ii) the extent to which our intent-to-treat estimates approximate average treatment effects given that we do not directly observe the age at which shoppers in our panel begin claiming public pension benefits. Point (i) also addresses the so-called "retirement consumption puzzle," whereby individuals fail to save enough for retirement and, consequently, observed proxies for consumption fall after

FIGURE F.3. Response of Store-Level Major Subcategory Average Prices to Payday



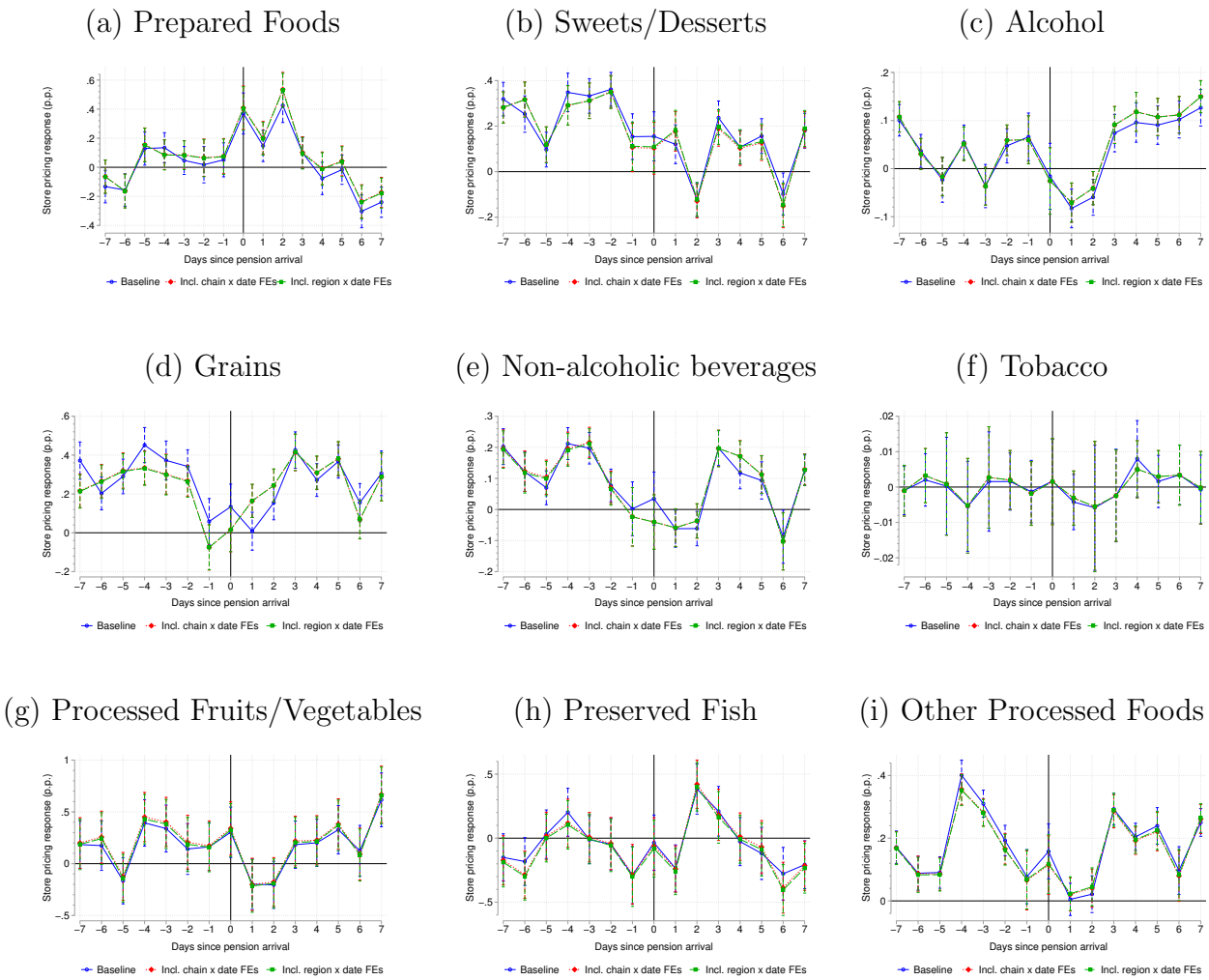
Notes: Each panel plots the event study coefficients $\widehat{\gamma}_j$ obtained from estimating regression (F.5) for spending within one of the one-digit goods categories defined in Appendix B. The y-axis records the percent increase in store-level prices around payday, using the average price index defined in (F.1). Point estimates in red obtained from augmenting (F.5) to include store chain $\times$ month-year fixed effects, while estimates in green result from including Census region $\times$ month-year fixed effects. Bars indicate 99% confidence intervals, with standard errors clustered at the store level.

FIGURE F.4. Store-Level Variety Responses within Each Major Goods Subcategory



Notes: Each panel plots the event study coefficients $\widehat{\gamma}_j$ obtained from estimating regression (5.2) where we use as the dependent variable the log of the daily number of goods sold (identified by unique barcodes), within one of the one-digit goods categories defined in Appendix B, and in a store relative to the average number sold over the sample time period. Point estimates in red obtained from augmenting (5.2) to include store chain × month-year fixed effects, while estimates in green result from including Census region × month-year fixed effects. Bars indicate 99% confidence intervals, with standard errors clustered at the store level.

FIGURE F.5. Store Pricing Responses around Payday by Major Goods Subcategory



Notes: We construct store-level price indices as in equation (F.1) and then estimate regression (F.5) to obtain estimates $\widehat{\gamma}_{1,j}$. We then construct a counterfactual last price index, using the last observed price for each barcode in the odd month preceding the payday event, and estimate (F.6) on the exact same sample using this price index as the dependent variable to obtain $\widehat{\gamma}_{2,j}$. Each panel in the figure repeats this procedure restricting to spending within one of the one-digit goods categories defined in Appendix B. The figure plots the point estimates and 99% confidence interval bars for the differences $\widehat{\gamma}_{1,j} - \widehat{\gamma}_{2,j}$, with standard errors in each regression clustered at the store level. Point estimates in red obtained from augmenting each regression with store chain $\times$ month-year fixed effects, while estimates in green result from including Census region $\times$ month-year fixed effects. The differences capture the percentage point increase in prices on payday that cannot be explained by changes in the composition of goods purchased.

retirees begin claiming benefits. This phenomenon is shown in recent studies by Agarwal, Pan, & Qian (2015) and Olafsson & Pagel (2020), using bank account and credit card data.

We run event study regressions where, instead of payday, the event is now defined as a shopper i passing the early or normal retirement age threshold y^*:

$$\frac{X_{i,c,my}}{\overline{X}_{i,my}} = \sum_{k=-12,k\neq-1}^{+12} \gamma_k \times \mathbb{1}\{age_{i,my+k} \geq y^*\} + \psi_{my} + \eta_i + \varepsilon_{i,c,my} \tag{F.8}$$

We define the forcing variable $age_{i,my}$ as the number of years passed as of month-year my since the month-year date of birth individuals report when they sign up for a loyalty point card with the retailer. To approximate true consumption responses, for this exercise, we define the expenditure ratio as the deviation from average monthly spending within the "raw foods" category of perishable goods defined in Appendix B.1. However, the spending paths are qualitatively similar even when we restrict to categories c which include semi-durable goods. Following Schmidheiny & Siegloch (2020), we bin the coefficients at the endpoints $k = -12$ and $k = +12$ to identify treatment effects since the staggered nature of individuals crossing either the early or normal retirement age threshold y^* renders the panel unbalanced in event-time.[17]

We estimate the event study coefficients γ_k by OLS and using the estimator of Sun & Abraham (2021), hereafter SA, which accounts for potential treatment effect heterogeneity induced by the fact that shoppers in our dataset will cross the retirement age threshold on different calendar dates depending on their date of birth. The SA estimator is applicable here because with treatment defined as $\mathbb{1}\{age_{i,my} \geq y^*\}$ we have a never treated group of individuals with $age_{i,my} < y^*, \forall my$.[18] In contrast, the naive OLS regression lumps together never-treated and not-yet-treated shoppers into the control group.

A potential issue with interpreting the coefficients γ_k in equation (F.8) as the effect of retirement on spending is that shoppers may have a spouse who retired before them. In such cases, we will estimate the effect of retirement after having a spouse who retired and potentially accumulated buffer stock savings out of benefit payments until the second spouse retired. One way to difference out the effect of intra-household transfers would be to interact the age cutoff dummy with an indicator $Paymonth$ equal to unity in months with scheduled

[17] In unreported results, we uncover similar event study spending patterns around retirement when we estimate a daily frequency version of (F.8) with the addition of day-of-week, week-of-month, and holiday fixed effects as in our baseline specification (4.1).

[18] In unreported results, we apply the estimator of de Chaisemartin & D'Haultfœuille (2020), which instead uses the not-yet-treated shoppers (i.e. those that reach retirement age within the panel) as a control group, but find these point estimates are similar to those obtained via the Sun & Abraham (2021) estimator. In our setting, the de Chaisemartin & D'Haultfœuille (2020) estimator is akin to a dynamic, fuzzy RD version of the event study design where we select a bandwidth around the retirement age cutoff to estimate treatment effects. An RD implementation would be inappropriate here to the extent that age is manipulated around the retirement cutoffs, since individuals may feel self-conscious about reporting their true age.

public pension paydays:

$$\frac{X_{i,c,my}}{\overline{X}_{i,my}} = \sum_{k=-12,k\neq-1}^{+12} \zeta_k \cdot \left(\mathbb{1}\{age_{i,my+k} \geq y^*\} \times Paymonth_{my} \right) + \psi_{my} + \eta_i + \varepsilon_{i,c,my} \quad \text{(F.9)}$$

The interaction term captures the additional bump in shoppers' spending on payday months after the shopper attains retirement age. The ζ_k yield the effect of retirement owing to the new income stream claimed by the spouse who recently became eligible for pension payments.

The results plotted in Figure F.5 indicate that allowing for shopper and spousal age cohort-specific heterogeneity are crucial to determining consumption trajectories following retirement. Regardless of whether we set $y^* = 60$ (Panel A) or $y^* = 65$ (Panel B), the OLS estimates display a fall in monthly spending on perishables around retirement, while the *SA* estimates show a bump in spending, at least within the first six months after retirement. The event studies represented by (F.9) in the right-hand side column of the figure show that in the pre-period, both the OLS and *SA* estimates may be confounded by the presence of previous spousal pension income. There is a sharp drop in spending during pension payment months leading up to retirement, perhaps due to the unobserved, already retired spouse using their pension payment income towards household spending. We lose statistical power in moving from equation (F.8) to (F.9), as there are only 22 payment months in our sample period.

F.5 Spending Responses by Age

Our baseline specification for documenting responses to pension receipt in equation (4.1) compares payday spending for shoppers above the normal retirement age (NRA) to those below the NRA. How does payday spending vary throughout the age distribution? Answering this question is relevant for evaluating the assumption we make in Section 6.4 that $x'(T)$ is similar for younger and older pensioners.

To this end, we run a pooled version of (4.1) which replaces $Payment_{i,t}$ using the NRA age cutoff with age bins interacted with the $Payday_t$ dummy:

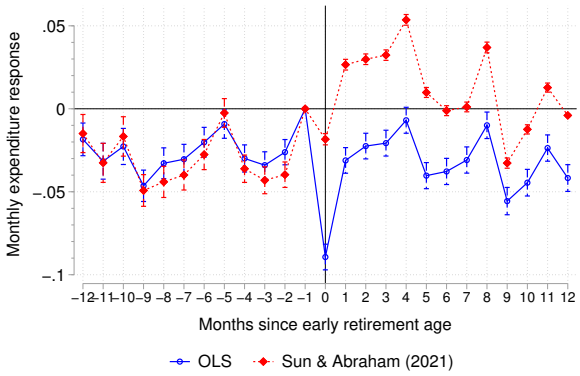
$$\frac{X_{i,t}}{\overline{X}_i} = \beta_b \cdot \mathbb{1}\{Agebin_i = b\} \times Payday_t + \delta_{dow} + \phi_{wom} + \varphi_{c,my} + \xi_h + \eta_i + \epsilon_{i,c,t} \quad \text{(F.10)}$$

where $\varphi_{c,my}$ is a set of chain by month-year fixed effects. Figure F.7 displays the results for our baseline sample of weekly shoppers (Panel A) and very frequent shoppers (Panel B), as defined in Appendix B.3. There is a small positive spending response on payday among the "control" group of shoppers below the early retirement age (ERA) 60 due to intra-household spillovers between eligible and non-eligible family members. As expected, payday spending is much higher among pension eligibles and matches our baseline responses in Table 1. There is no statistically significant difference between the spending responses among cohorts above the ERA. Moreover, such differences are smaller for very frequent shoppers – who visit a store every other day – which helps eliminate concerns about transactions becoming more selected at higher age bins due to mobility constraints among older retirees. These results

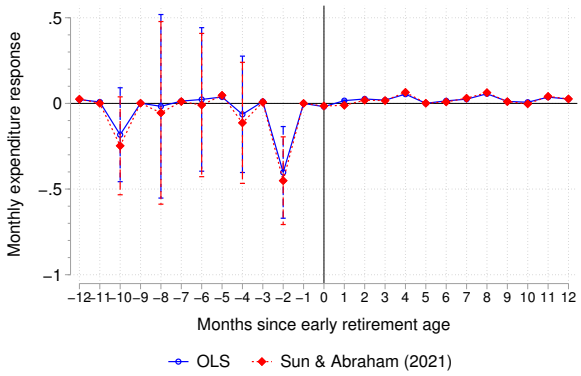
FIGURE F.6. Spending Responses around Early and Normal Retirement Age

A. Early Retirement Age ($y^* = 60$ years old)

$\mathbb{1}\{age \geq y^*\}$

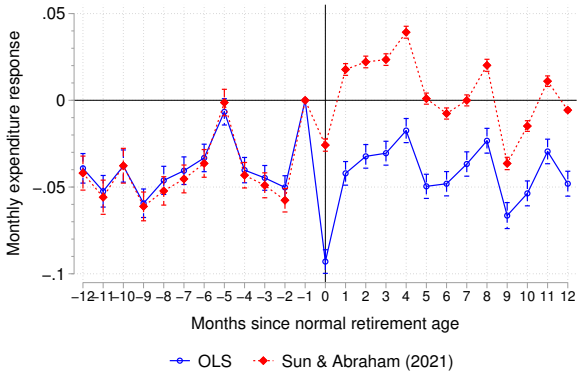


$\mathbb{1}\{age \geq y^* \times Paymonth\}$

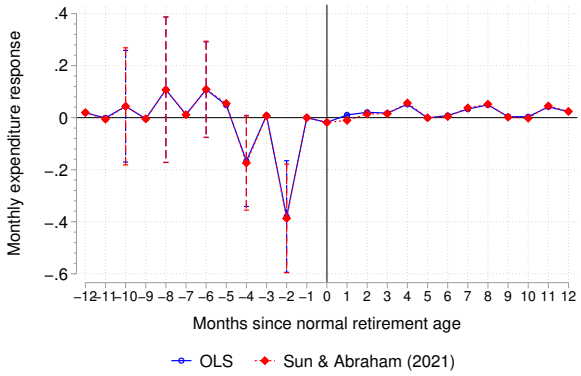


B. Normal Retirement Age ($y^* = 65$ years old)

$\mathbb{1}\{age \geq y^*\}$

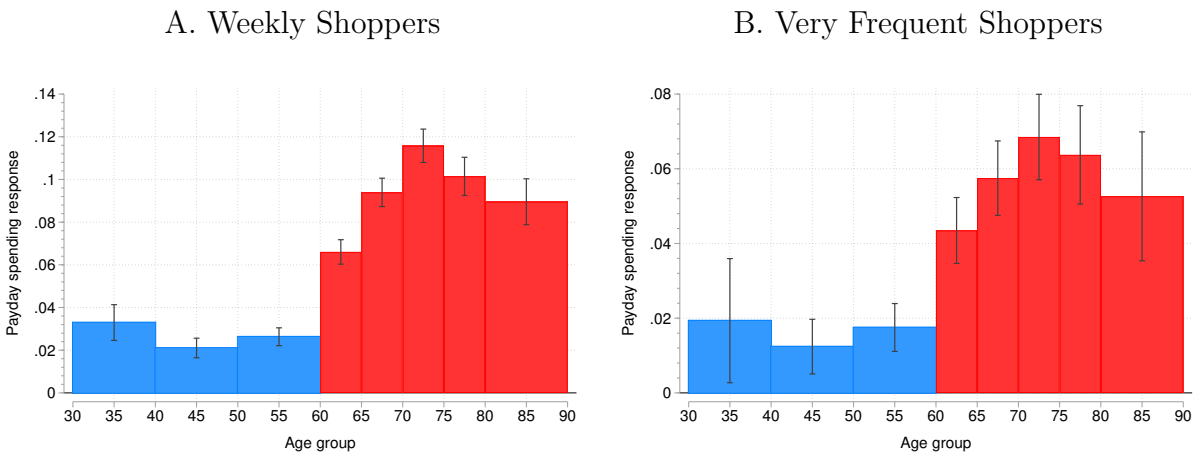


$\mathbb{1}\{age \geq y^* \times Paymonth\}$



Notes: Left-hand side panels plot the coefficients γ_k estimated from event study equation (F.8) by either OLS or the Sun & Abraham (2021) estimator. Right-hand side panels instead plot the coefficients ζ_k from equation (F.9), which isolates the effect of shoppers newly receiving income from the pension system. Panel A performs this exercise using shoppers crossing the early retirement age threshold (60 years) while Panel B does the same using the normal retirement age (65 years) as the cutoff. We bin coefficients at the endpoints $k = -12$ and $k = +12$ so that those point estimates represent the effects at 12 months or longer after or before retirement. All point estimates are relative to the response one month prior to retirement ($k = -1$). The y-axis represents the percent increase in monthly spending on perishables ("raw foods") relative to average monthly spending in that category over the panel. As in our baseline results, we restrict to a fixed panel of weekly shoppers. We winsorize the top 1% of monthly expenditures. Bars indicate 99% confidence intervals, with standard errors obtained from clustering by shopper ID.

FIGURE F.7. Shoppers' Payday Responses (Difference-in-Differences) by Age Bin



Notes: We estimate pooled versions of our baseline difference-in-differences specification (4.1) with age bins as defined by (F.10). All specifications are on the intensive margin with a full set of time and chain fixed effects. Panel A does this for a fixed panel of either weekly shoppers, and Panel B for very frequent shoppers for whom we can match average monthly grocery spending in the nationally representative FIES (see Appendix B.3). The age bins are: 20-29 (omitted category), 30-39, 40-49, 50-59, 60-64, 65-69, 70-74, 75-79, 80-89. Ages are censored below 20 and above 90 years old in our data. Blue coefficient bars indicate age bins below the early retirement age of 60, while red coefficient bars indicate spending responses for age bins eligible for claiming pension payments. Capped bars indicate the 99% robust confidence intervals with standard errors clustered by shopper ID.

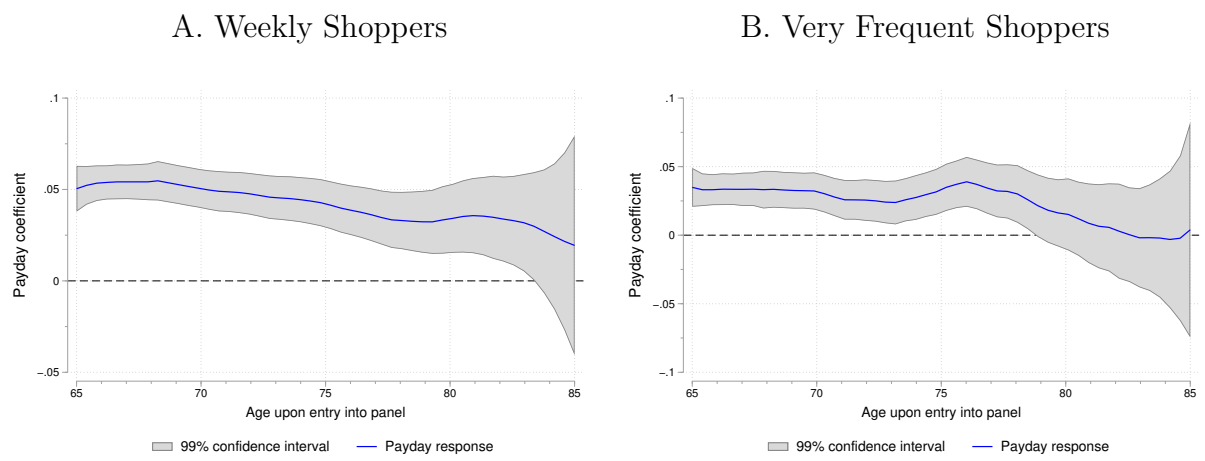
validate the assumption in our counterfactual exercises in Section 6.4 in which we compare costs of changing the retirement eligibility age vs. altering the payment frequencies.

Figure F.8 shows how payday spending varies continuously with respect to age for shoppers who enter the panel above the NRA. For this exercise, we estimate a single-differenced regression comparing payday vs. non-payday spending within each shopper's transaction history. The results complement those in Figure F.7; there is a slight negative gradient between the β^i and age, but this only manifests in the very frequent shopper sample starting at ages in the late seventies, and the confidence intervals become wider at that point.

F.6 Spending Responses by Store Quality

Are the spending responses shoppers exhibit on pension paydays due to liquidity constraints (i.e. being unable borrow against future payments or dip into savings) or internality problems? One way to indirectly test for the presence of liquidity constraints is to rank shoppers on the basis of the quality of the stores they frequently visit. This exercise builds on the insights of Bils & Klenow (2001), who show that quality Engel curves slope upward – that is, as income rises consumers spend more on high-quality goods. Rather than taking a stance on which barcodes represent high vs. low-quality goods within each category, we take a revealed preference approach and assume that higher-quality goods command higher

FIGURE F.8. Retirees' Payday Responses as a Function of Age



Notes: We estimate the time series regression $X_{i,t}/\overline{X}_i = \beta^i \cdot Payday_t + \delta_{dow} + \phi_{wom} + \psi_{my} + \xi_h + \epsilon_t^i$ for each individual shopper ID using raw foods spending. The figure fits a local linear function to the relationship between payday responses $\widehat{\beta^i}$ and the shopper's age as of when they first entered the panel. Panel A does this for a fixed panel of either weekly shoppers, and Panel B for very frequent shoppers for whom we can match average monthly grocery spending in the nationally representative FIES (see Appendix B.3). We winsorize the top 1% of daily expenditures. 99% confidence intervals represented by the gray shaded area.

prices. Stores or chains which display greater (weighted) average prices of goods sold are then selling higher-quality goods than their competitors.[19]

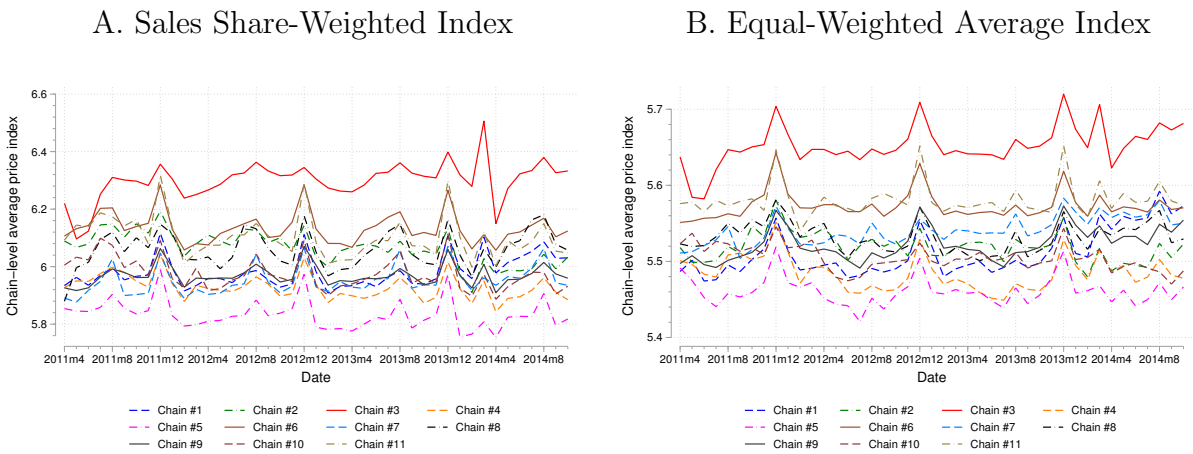
Our store quality measure averages over daily store-level prices:

$$\widetilde{\Phi}_s = \frac{1}{|T^{np}|} \sum_{t \in T^{np}} \Phi_{s,t} = \frac{1}{|T^{np}|} \sum_{t \in T^{np}} \left(\sum_{k \in \Omega_{s,t}} \omega_{k,s,t} \log p_{k,s,t} \right) \tag{F.11}$$

where we average over days in the set T^{np} outside a $\pm$ 3-day window around pension paydays to avoid reverse causality due to pensioners' substituting towards higher-quality goods and stores on paydays, as documented in Section 5.2. As before, $\Omega_{s,t}$ is the set of barcodes k transacted in that store on a given date. $\omega_{k,s,t}$ is the weight of each good k in the store's price index. Previously, we took equal-weighted averages, or $\omega_{k,s,t} = 1/n_{s,t}$, with $n_{s,t}$ the number of barcodes transacted in store s on t. Our store quality rankings are strongly correlated regardless of whether we use an equal-weighted price index, daily sales shares, or time-invariant sales shares to weight products in a store's inventory. Versions of the store quality index residualized on the store's geographic region are 90% correlated with the unresidualized index, suggesting that our measure of store quality is not simply picking up cost of living differences across locations. To produce a chain-level quality ranking, we

[19] In principle, it is possible that using average prices as a quality rank picks up differences in market power across retailers. This is fine for our purposes, since higher-income shoppers will tend to be more demand inelastic and sort into stores which offer fewer temporary sales. Applying the temporary sales filters from Appendix D to extract regular prices barely changes our store quality ranking.

FIGURE F.9. Quality Index of Retail Chains Using Price of Goods Sold



Notes: The figure plots the monthly chain-level version of equation (F.11) for the top 11 chains in our sample which very frequent shoppers visit. That is, we compute $\widetilde{\Phi}_s$ for each store-month and then aggregate within the chain by taking a sales-weighted average across stores in the chain. Panel A displays the chain-level indices using daily store-level sales share weights for $\omega_{k,s,t}$, while Panel B sets $\omega_{k,s,t} = 1/n_{s,t}$, so that each product k within a story on a day receives equal weight.

take a sales share-weighted average of $\widetilde{\Phi}_s$ across all stores in the chain.

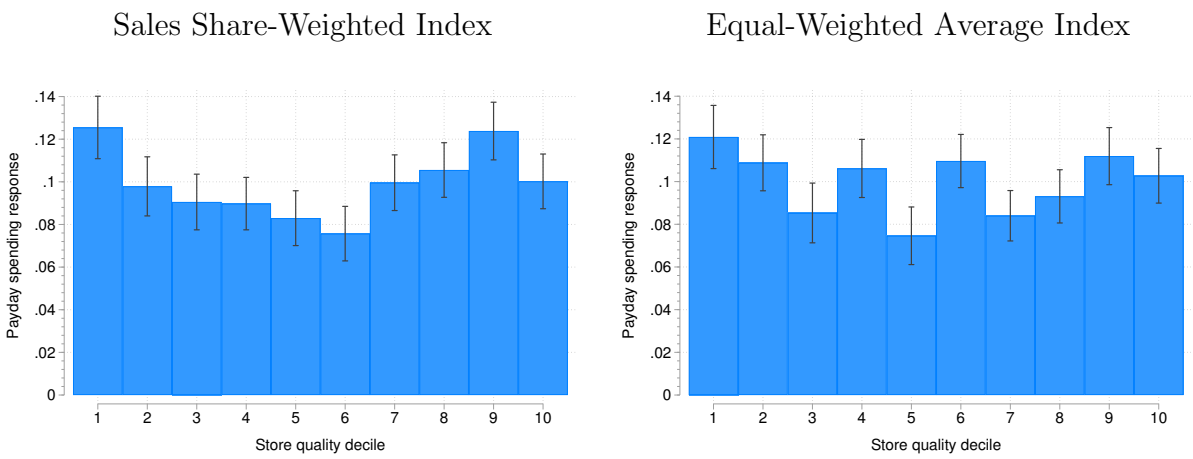
As a proof of concept, Figure F.9 plots the monthly chain-level pricing indices for the 11 largest chains in our sample which are those visited by the very frequent shoppers whose spending levels are nationally representative. While there is common seasonality across chains, equation (F.11) produces clear level differences. For example, over the full time sample, the highest-quality chain (#3) charges a share-weighted average price which is 8.0% higher than the lowest-quality chain (#5) and 5.3% higher than the median chain.

The average shopper in our sample visits two unique stores (the median is one) and stays within one chain during their time in the panel. 50% of shoppers always visit the same store, and 91% always shop within the same chain. For shoppers who switch, we try two methods to assign them to a store quality measure. One is to simply use the $\widetilde{\Phi}_s$ corresponding to each shopper's store that they visit the most frequently. Alternatively, we define for each shopper an expenditure share-weighted average of $\widetilde{\Phi}_s$, where the shares are computed to exclude spending outside the $\pm$ 3-day window around paydays. The results are very similar regardless, and so we simply assign each shopper to the quality index of their revealed-preferred store.

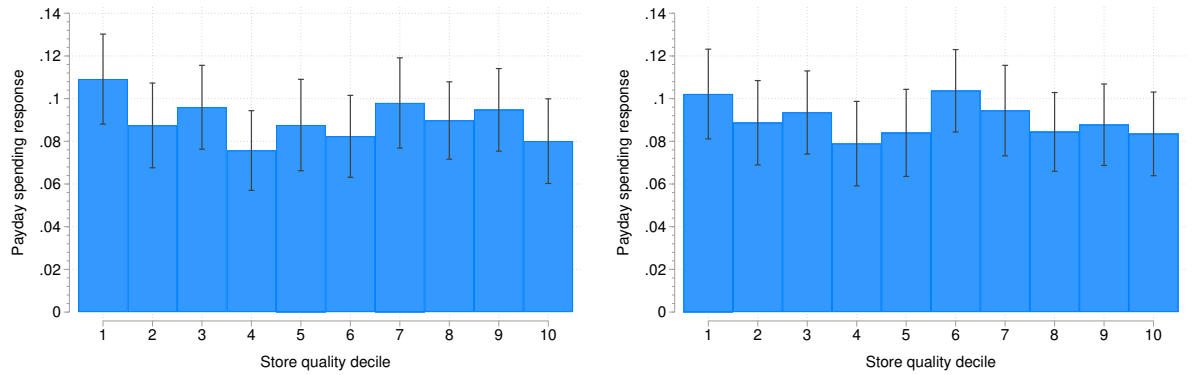
Figure F.10 shows that there is no statistically discernible, monotonic relationship between store quality and the degree to which shoppers deviate from their consumption path on pension paydays. Spending responses are even more uniform after we remove any geographic variation in cost of living by ranking shoppers on quality of stores within the Tokyo metropolitan area (Panel B). Since shoppers' selected store quality rises with their disposable income, we would expect there to be a negative gradient if liquidity constraints played a prominent role. We therefore interpret these results as inconsistent with a liquidity constraint explanation for spikes in spending on payday.

FIGURE F.10. Shoppers' Payday Responses by Favorite Store Quality Index Decile

A. Using Stores in All Census Regions



B. Using Only Stores in the Tokyo Metropolitan Area



Notes: We estimate pooled versions of our baseline difference-in-differences specification (4.1) for each decile of the store quality index defined by (F.11). We match shoppers to a store quality measure on the basis of their most-visited store. All specifications are on the intensive margin with a full set of time fixed effects. Panel A uses all stores, and Panel B takes out the region fixed effect by repeating the analysis including only stores located in the Tokyo metropolitan area, with deciles of $\widetilde{\Phi}_s$ re-calculated for stores within Tokyo. Capped bars indicate the 99% robust confidence intervals with standard errors clustered by shopper ID.

Appendix References

Abe, N. & A. Tonogi (2010): "Micro and Macro Price Dynamics in Daily Data," *Journal of Monetary Economics*, 57(6): 716-728.

Agarwal, S., J. Pan, & W. Qian (2015): "The Composition Effect of Consumption around Retirement: Evidence from Singapore," *American Economic Review: Papers & Proceedings*, 105(5): 426431.

Aguiar, M. & M. Bils (2015): "Has Consumption Inequality Mirrored Income Inequality?" *American Economic Review*, 105(9): 2725-2756.

Aguiar, M. & E. Hurst (2007): "Life-Cycle Prices and Production," *American Economic Review*, 97(5): 1533-1559.

Baugh, B. & J. Wang (2018): "When Is it Hard to Make Ends Meet?, *mimeo*, University of Illinois at Urbana-Champaign.

Bils, M. & P.J. Klenow (2001): "Quantifying Quality Growth," *American Economic Review*, 91(4): 1006-1030.

Bils, M. & P.J. Klenow (2004): "Some Evidence on the Importance of Sticky Prices," *Journal of Political Economy*, 112(5): 947-985.

Blinder, A.S., E.R.D. Canetti, D.E. Lebow, & J.B. Rudd (1998): *Asking About Prices: A New Approach to Understanding Price Stickiness*, Russell Sage Foundation: New York.

Bryan, G., D. Karlan, & S. Nelson (2010): "Commitment Devices," *Annual Review of Economics*, 2(1): 671-698.

Chahrour, R.A. (2011): "Sales and Price Spikes in Retail Scanner Data," *Economics Letters*, 110(2): 143-146.

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

Eichenbaum, M., N. Jaimovich, S. Rebelo (2011): "Reference Prices, Costs, and Nominal Rigidities," *American Economic Review*, 101(1): 234-262.

Kehoe, P.J. & V. Midrigan (2008): "Temporary Price Changes and the Real Effects of Monetary Policy," NBER Working Paper, No. 14392.

Kehoe, P.J. & V. Midrigan (2015): "Prices Are Sticky After All," *Journal of Monetary Economics*, 75: 35-53.

Kondo, K. (2019): "Municipality-level Panel Data and Municipal Mergers in Japan," RIETI Data Management Project, https://www.rieti.go.jp/en/publications/summary/19030013.html.

Kueng, L. (2018): "Excess Sensitivity of High-Income Consumers," *Quarterly Journal of Economics*, 133(4): 1693-1751.

Nakamura, E. & J. Steinsson (2008): "Five Facts About Prices: A Reevaluation of Menu Cost Models: Supplementary Material," *Quarterly Journal of Economics*, 123(4): 1415-1464.

Olafsson, A. & M. Pagel (2020): "Repaying Consumer Debt and Increasing Savings After Retirement," *mimeo*, Columbia Business School.

Schmidheiny, K. & S. Siegloch (2020): "On Event Studies and Distributed-Lags in Two-Way Fixed Effects Models: Identification, Equivalence, and Generalization," CEPR Discussion Paper, No. 13477.

Sudo, N., K. Ueda, K. Watanabe, & T. Watanabe (2018): "Working Less and Bargain Hunting More: Macroimplications of Sales during Japan's Lost Decades," *Journal of Money, Credit and Banking*, 50(2-3): 449-478.

Sun, L. & S. Abraham (2021): "Estimating Dynamic Treatment Effects in Event Studies with Heterogeneous Treatment Effects," *Journal of Econometrics*, 225(2): 175-199.

Unayama, T. (2018): *Introduction to Japanese Household Surveys*, Springer: Tokyo.

Vellekoop, N. (2018): "Explaining Intra-Monthly Consumption Patterns: The Timing of Income or the Timing of Consumption Commitments?" *Leibniz Institute for Financial Research (SAFE)* Working Paper.

Zhang, C.Y. (2023): "Consumption Responses to Pay Frequency: Evidence from 'Extra' Paychecks," *mimeo*, University of Wisconsin-Madison.