Consumer Inventory and the Cost of Living Index: Theory and Some Evidence from Japan

Kozo Ueda* Kota Watanabe† Tsutomu Watanabe‡

December 1, 2019

Abstract

This paper examines the implications of consumer inventory for the cost of living index and business cycles. We begin by providing stylized facts about consumer inventory using scanner data. We then construct a quasi-dynamic model to describe consumers' purchase, consumption, and inventory behavior. A key feature of our model is that inventory is held by warehouse firms, not by consumers, which enables us to continue to construct the cost-of-living index in a static manner even in an economy with storable goods. Using this model, we show that stockpiling during temporary sales generates a huge chain drift in the conventional cost of living indexes, which are constructed without paying attention to consumer inventory. The chain drift of the cost of living index derived in our model is greatly mitigated since it is based on consumption prices (rather than purchase prices) and quantities consumed (rather than quantities purchased). We provide empirical evidence supporting these theoretical predictions. We also show empirically that consumers' inventory behavior tends to depend on labor market conditions.

JEL Classification Number: C43, D15, E31

Keywords: inventory; cost-of-living index; price elasticity

*Waseda University (E-mail: kozo.ueda@waseda.jp)

†Canon Institute for Global Studies and University of Tokyo (E-mail: watanabe.kota@canon-igs.org)

‡University of Tokyo (E-mail: watanabe@e.u-tokyo.ac.jp). We thank conference and seminar participants at the Keio-Waseda Workshop, the Ottawa Group Meeting 2019 (Rio), and the UBC for useful comments and suggestions. This research is funded by the JSPS Grant-in-Aid for Scientific Research (16KK0065 for Ueda and 18H05217 for T. Watanabe). All remaining errors are our own.

1 Introduction

Many goods (even vegetable and milk) are storable, although most economic models deal with perishable goods, neglecting intertemporal substitution. This implies that purchase (which is often observable) is not necessarily equal to consumption (which is often unobservable), and their difference serves as consumer inventory. In particular, temporary sales and the anticipation of an increase in a value-added tax rate often increase purchase more than consumption. This stockpiling behavior of consumers poses challenges to economists, for example, in the measurement of price index. It is well known that the Törnqvist index is a good approximation of the cost-of-living index (COLI) (Diewert 1976). However, as shoin in Figure 1, the Törnqvist index has a huge downward bias, so called, the chain drift, when we apply the formula to Japan's scanner data at a daily basis (details will be explained later). The price level has decreased to 10^{-10} in a quarter century by experiencing 60% deflation annually. Furthermore, stockpiling may matter for the measurement of price elasticity (Hendel and Nevo (2004) and Cashin and Unayama (2016)) as well as the macroeconomy.

In this study, we examine consumers' stockpiling behavior in the following way. First, we document stylized facts associated with storable goods by employing both retailer and home scanner data for Japan. Note that the latter data are novel in that they record when households purchase, when they consume, and when they consume all their inventory, at a household and product level on a daily frequency. We report three stylized facts: (1) the chained price index entails a huge bias, where the Paasche and Törnqvist index have deflationary and the Laspeyres index is inflationary; (2) there is a significant difference in the amount of purchase before, during, and after a sales period; (3) consumption is dependent on inventory.

Second, we construct a simple model of storable goods to explain facts, infer consumption, and construct price index. The model incorporates stockpiling behavior in an environment where prices take either high or low values exogenously. The quantity purchased increases at sales, part of which is for a future use after price becomes high. The novel feature of the model is that it is quasi dynamic in that warehouse firms hold inventory instead of households and solve a dynamic optimization problem, while households do not. This helps simplify the calculation of the COLI, with many key properties being kept, because the household's cost minimization problem is static. This is an

advantage over so-called dynamic COLI that is often complex and hard to interpret.[1] Using the model, we explain the three stylized facts mentioned above. In particular, we show that the chain drift is attributed to asymmetry due to the non-negative constraint of inventory and purchase.

Based on the model, we propose a new method to infer consumption and inventory only by using the retailer scanner data, and in turn, to calculating the COLI. We show that consumers' stockpiling behavior can be conveniently summarized by a single variable: the degree of stockpiling at sales, which expresses how long an inventory lasts after sales end. Applying our method to Japan's retailer scanner data, we confirm that it explains the bias in the chained price index well. A newly proposed price index based on consumption weight mitigates a downward chain drift of the Törnqvist index considerably, but not perfectly.

Finally, we investigate macroeconomic implications of goods being storable. Although macroeconomists often neglect temporary sales, this cannot be justified if households' stockpiling at sales depends on business cycles and revenue coming from sales is big. Indeed, the latter has reached 30% of total sales during the 2000s in Japan (Sudo et al. 2018). Furthermore, the anticipated increases in the consumption tax rate in April 1997 and April 2014 caused big stockpiling in March and the falloff in the following several months. We further find that, when the unemployment rate is low or hours worked are long, the degree of stockpiling tends to be large. This implies that an increase in the amount of sales and the quantity purchased in booms is magnified by intensified stockpiling, while an increase in the unit price is mitigated.[2]

There are at least four strands of research related to our study. The first strand concerns household inventory of storable goods. Boizot, Robin, and Visser (2001), Griffith et al. (2009), Hendel and Nevo (2006a), and Kano (2018) document the pattern of household purchase and consumption empirically. Among them, Boizot, Robin, and Visser (2001) and Kano (2018) are noteworthy in the rare use of household inventory data like us. In theory, our model is the closest to that developed by Hendel and Nevo (2006a).

[1]For example, Ueda (2019) shows that an individual price decrease (increase) causes the dynamic COLI to increase (decrease) because a household spends a lot for stockpiling when price is low and spends none when price is high. Even if we use alternative measure of the dynamic COLI, it does not satisfy the circularity test.

[2]Our model indicates that the degree of stockpiling is endogenous, and that it is closely related to the size of sales discounts and the frequency of sales. However, we empirically show that, even after controlling for the sale channel, the degree of stockpiling depends on business cycles.

Compared with these studies, the novelty of our study is in the focus on the price index (COLI) and macroeconomic implications. For this objective, our model is original in that it is quasi dynamic and separates households into households in a narrow sense and warehouse firms who can hold inventory. It should also be noted that, in this strand of literature, consumption is often assumed to be constant (e.g., Boizot, Robin, and Visser (2001) and the empirical part of Hendel and Nevo (2006a)), although Hendel and Nevo (2006a) and Kano (2018) emphasize the possibility that consumption is state dependent.[3] Our approach takes state-dependent consumption into serious account, which is confirmed in our data and then used for constructing the COLI.

Literature on the chain drift or the time aggregation problem in the chain index includes Frisch (1936), Reinsdorf (1999), Feenstra and Shapiro (2003), ILO et al. (2004), and Ivancic, Diewert, and Fox (2011). In the literature, it seems a concensus to attribute the chain drift to stockpiling (sometimes called, quantity bouncing). To this problem, Feenstra and Shapiro (2003) propose the use of virtually fixed base index, while Ivancic, Diewert, and Fox (2011) and Haan and van der Grient (2011) propose the so-called GEKS. The former approach is simple but presumes perfect foresight in calculating weights and discards data on short-lived or new goods. The latter approach is constructed to satisfy the circularity test, which is one of key properties for price index. However, a caveat is its complexity. Also, it is subject to revisions each time new data become available. Compared with these studies, our study is novel in attempting to infer varying consumption from retailer scanner data, which is important for not only economists but also practitioners who compile price index.

While studies on the COLI date back to Könus (1924), recent contributions include Feenstra and Shapiro (2003), Chevalier and Kashyap (2018), Reis (2009), Gowrisankaran and Rysman (2012), Wang (2013), Osborne (2018), and Ueda (2019). The first two papers propose a proxy for the COLI in the presence of sales, although the problem to compute the COLI remains static. The other papers study dyanmic COLI, which encompasses intertemporal substitutionis by households. In paticular, Osborne (2018) is most closely related. He incorporates both storable goods and sales and computes the dynamic COLI by calculating the sequence of tax or subsidy, which makes the period utility constant over time. While this is indeed a big contribution, interpreting the short-run movements of the dynamic COLI is often hard. Because our model is quasi

[3]See also Hendel and Nevo (2004), Cashin and Unayama (2016), Erdem, Imai, and Keane (2003), Hendel and Nevo (2006b), and Osborne (2018).

dynamic and households' optimization problem is static, it becomes far simpler to define, interpret, and construct the COLI, which we believe is beneficial to practitioners as well as economists.[4]

Last, there is a vast literature on temporary sales. Related to storable goods are a survey by Hendel and Nevo (2004), empirical studies by Blattberg and Neslin (1989), Neslin and Schneider Stone (1996), and Hendel and Nevo (2003), and theoretical studies by Salop and Stiglitz (1982), Hong, McAfee, and Nayyar (2002), Hendel, Lizzeri, and Roketskiy (2014). The empirical studies are analyzed from a marketing perspective. The thoretical studies are from a firm side, constructing a model to explain why firms set dispersed prices for the same goods, while our study is a household-side analysis. It should also be noted that, compared with these papers, our study neglects household heterogeneity to simplify our analysis and concentrate on macroeconomic implications.

The rest of the paper is organized as follows. Section 2 provides stylized facts by using the two kinds of scanner data in Japan. Section 3 develops a quasi dynamic model to embed stockpiling behavior. Section 4 extends the model to infer consumption from retailer scanner data and apply this. Section 5 concludes.

2 Facts from Japan's Scanner Data

2.1 Two Scanner Data

We use two kinds of scanner data in Japan. The first one is from a retailer side: the Point-of-sales (POS) scanner data collected by Nikkei Inc. The data record the number of units sold and the amount of sales (price times the number of units sold) for each product i and retail shop s on a daily basis t. The observation period runs from March 1, 1988 to September 30, 2013. Products recorded include processed food and daily necessities, covering 170 of the 588 categories in the Consumer Price Index (CPI) and making up about 20 percent of households' expenditure. See Abe and Tonogi (2010), Sudo, Ueda, and Watanabe (2014), and Sudo et al. (2018) for detailed data properties.

[4]Osborne (2018) finds that the price indexes proposed by Feenstra and Shapiro (2003) and Chevalier and Kashyap (2018) approximate the dynamic COLI better than the fixed-base indexes such as Laspeyres. However, his analysis is limited to particular products (canned tuna and canned soup) for particular periods, so there is a possibility that this finding is just a coincidence. Since the price index we propose in this study is the COLI, we do not need to search for alternative measure to approximate the COLI.

When there is no record on the POS scanner data for product i,shop s, and date t, neither quantity nor price can be observed, and thus, early studies usually omit such a sample. However, this seems to occur often after sales, implying that the stockpiling during sales lead to zero purchase after sales end. Because this conveys important information on consumer inventory, in this study, we interpolate such a sample. Specifically, we set the quantity purchased at zero and the price at the regular price on the nearest past date.[5]

The other data are from a household side: Shoku-map (literal translation is food-map) scanner data collected by Lifescape Marketing Co. Respondents are mainly housewives, whose number is about 400 in each period (about 4,000 in total). The data record the number of units purchased and the date of consumption for each product i and household j on a daily basis t. Moreover, they record when consumption ends (used up, wasted, etc.) for each i and j. The observation period runs from September 1, 1998 to XXX. Note that products recorded are food only and there is no information on the price purchased. Another limitation is that there is no information on how much (i.e. weights) product i is consumed in each time of consumption. The data record the number of units both purchased and consumed, which is sufficiently useful if products are consumed discretely such as a cup of instant noodle and a can of beer. However, for products like salt, we cannot know by how much household j uses product i, although we can know the date she uses it.[6] For example, Figure 2 shows a consumption pattern of a certain salt i for certain household j, where each vertial line represents a consmption flag. This displays that she purchased salt on date $t = 19$, started using from $t = 22$, and finished using on $t = 144$. The inventory duration is thus $144 - 19 + 1 = 126$ days.

In both of the data, all products are identified by the Japanese Article Number (JAN) code, which enables us to merge them. Further, we classify products into groups using the 3-digit product categories provided by Nikkei Inc. There are 214 categories in total, such as instant cup noodles, yogurt, beer, and toothbrushes.

Figure 3 illustrates purchase and consumption patterns for the 3-digit product category of beer and a particular household using the Shoku-map data. The horizontal axis represents dates, while the vertical axis represents the cumulative number of purchased products. The left and right ends (in green) of each horizontal line show when products

[5]We do not interpolate samples for the right-hand panel of Figure 1 and Figure 5.

[6]If she uses salt i N times in a day, the data record N, where N is an integer greater than or equal to zero.

are purchased and when they are consumed, respectively. This household seems to purchase more than one bottles of beer (six or a dozen) around when she consumes up all her inventory, while she consumes beer almost every day.

Figure 4 shows the density of consumption durations using the Shoku-map data. We count the durations among the date of puchase (t_p), the date of the first consumption (t_f), and the date of the last consumption (t_l), for each product i and household j. Regarding $t_l - t_p$, the figure shows that about 8% are consumed all on the same date of purchase. The mean duration of $t_l - t_p$ is 28 days, while the mean of $t_f - t_p$ and $t_l - t_f$ is both about 14 days.

In the following analyses using the POS data, we identify temporary sales by employing the V-shape filter. More concretely, we follow the procedure explained in Nakamura and Steinsson (2010) using their sales filter A with a window of $L = K = J = 42$ days.[7] Observations for product k on day t are classified as sales if and only if price p_t^k deviates from its regular price $\overline{p}_t^k$ by more than two yen.

Futhermore, it needs to be emphasized that, unlike most previous studies including ours, we interpolate missing observations by setting $p_t^k = \overline{p}_t^k$ and $x_t^k = 0$. By the nature of the POS data, no observation is recorded if product k is not purchased on day t. Because this often occurs between the periods when prices increase from sales to regular ones, we suspect that this reflects stockpiling, which is central to our study. Thus, we decide not to neglect missing observations and interpolate them by regular price $p_t^k = \overline{p}_t^k$ and zero purchase quantity $x_t^k = 0$, if we have observations after t (i.e., unless the product exits from the market permanently).

2.2 Three Facts

Using the two data, we document three stylized facts, which are closely related to goods storability.

[7] The V-shape filter does not mean that sales end in just one day. See Sudo et al. (2018) for robustness analyses regarding the choice of sales filters.

Fact 1. When weights are based on purchase, the changes in the price indexes have the following inequality: $\pi^P < \pi^T < 0 < \pi^L$.

Let us denote the price and the quantity of product k in period t by p_t^k and x_t^k. Changes in the chained price indexes from $t - dt$ to t, π_t^X $(X = L, P, T)$, are defined by[8]

$$\pi_t^L = \sum_{k \in K_{t-dt} \cap K_t} W_{t-dt}^k (K_{t-dt} \cap K_t) \log \left(\frac{p_t^k}{p_{t-dt}^k} \right), \tag{1}$$

$$\pi_t^P = \sum_{k \in K_{t-dt} \cap K_t} W_t^k (K_{t-dt} \cap K_t) \log \left(\frac{p_t^k}{p_{t-dt}^k} \right), \tag{2}$$

$$\pi_t^T = \sum_{k \in K_{t-dt} \cap K_t} \frac{W_{t-dt}^k (K_{t-dt} \cap K_t) + W_t^k (K_{t-dt} \cap K_t)}{2} \log \left(\frac{p_t^k}{p_{t-dt}^k} \right), \tag{3}$$

respectively, when they are based on Laspeyres, Paasche, and Törnqvist.[9] The weight share $W_t^k(K_{t-dt} \cap K_t)$ equals $p_t^k x_t^k / \sum_{k' \in K_{t-dt} \cap K_t} p_t^{k'} x_t^{k'}$, where $k \in K_{t-dt} \cap K_t$ represents the matched-sample product that exist both in $t - dt$ and t. When $dt = 1$, we can construct the chained price index by the cumulative sum of the past price changes: $P_t^X = \exp\left(\sum_{s=1}^t \pi_s^X\right)$ for $X = L, P, T$.

It should be emphasized that the choices of p_t^k and x_t^k play a very important role. For the time being, we use conventional p_t^k and x_t^k that are observable from the POS data. That is, p_t^k and x_t^k represent the price purchased and quantity purchased, respectively. Later, we try to infer consumption-based p_t^k and x_t^k. To distinguish from the conventional ones, we express them by r_t^k and c_t^k,respectively, and the change in the price index by π_t^*.

Using the POS data, we calculate the time-series of the price level (P_t^X), where we normalize the innitial price level as one. The results are striking. The left-hand panel of Figure 1 shows that the price level increases by almost 10^{72} in a quarter century when it is based on Laspeyres, while it decreases to almost 10^{-90} when it is based on Paasche. The size of price changes is milder, but continues to big when it is based on Törnqvist. The price level decreases to almost 10^{-10}. Thus, the following inequality holds: $\pi^P < \pi^T < 0 < \pi^L$, exhibiting a downward bias for Paasche and Törnqvist and an upward bias for Laspeyres. In the right-hand panel, we draw the average price changes in

[8] In this study, we focus on the chained price index, not on the fixed-base price index. See ILO et al. (2004) and Ivancic, Diewert, and Fox (2011) for the discussions on the two.

[9] There are various types of Laspeyres and Paasche indexes. The ones we use in this study are called logarithmic Laspeyres and logarithmic Paasche.

a quarter century based on Törnqvist, where we employ different time intervals dt from 1 day to 365. This shows that the average deflation rate is about 60 percent annually if $dt = 1$, but such a downward bias becomes smaller as dt increases. When $dt = 365$, each price is compared with the one on the same date in the previous year. Although the average inflation rate continues to be negative, it is just about -0.3 percent annually. As Figure 5 shows, its time-series developments becomes very similar to those in the changes in the Consumer Price Index (CPI) for grocery that represents the CPI of the same product category as the POS data.

This big bias, both in downward and upward directions, arises from quantity bouncing or stockpiling. To see this, let us consider a simple case of sales as is shown in Table 1.[10] There are three periods: $t = 1$, 2, and 3, and two products: A and B. In period $t = 2$,the price of product A drops temporarily (sales). In such an enviornment and if the elasticity of substitutions is greater than one, it is natural to observe a higher share of product A in period $t = 2$ than in period $t = 1$ ($W_1 < W_2$). Furthermore, because consumers stockpile (that is, goods are storable), a share of product A in period $t = 3$ should be smaller than in period $t = 1$ ($W_3 < W_1$). That is, households have enough inventory on a day after sales, so they do not need to buy as much as they do before sales. Then, it is immediate to see $\pi^P < \pi^T < 0 < \pi^L$ from equations (1) to (3).[11]

Fact 2. The quantity purchased just before sales tends to be greater than that purchased just after sales. The quantity purchased during the first half of sales tends to be greater than that purchased during the second half of sales.

We examine whether the pattern shown in Table 1 is indeed the case by using the POS data. To do this, we collect the following variables for each goods k and sales event s: P_H and X_H^1 denote the price and the quantity purchased just before sales, respectively; P_L^1 and X_L^1 denote the average price and quantity purchased at the first half spell of

[10] Similar discussions have been already made. See, for example, Haan and van der Grient (2011).

[11] In an extreme case, product A is not purchased at all in period $t = 3$, that is, $W_3 = 0$. This causes the missing observation problem. If we do not interpolate this by setting $p_{t=3}^A = \overline{p}_t^A$ and $x_t^A = 0$, a price increase is fully neglected from $t = 2$ to 3 because product A is not in the matched sample of $k \in K_{t=2} \cap K_{t=3}$. Furthermore, even if the share of product A picks up at $t = 4$, no price increase is embedded at that time, either, because product A is not in the matched sample of $k \in K_{t=3} \cap K_{t=4}$. Therefore, our interpolation is considered to make the downward bias of the Törnqvist index milder.

sales, respectively; P_L^2 and X_L^2 denote the average price and quantity purchased at the second half spell of sales, respectively; P_H and X_H^2 denote the price and the quantity purchased just after sales, respectively; and T denotes the spell of sales. Note that we identify sales events when a price after sales returns the price before sales (i.e., P_H).

We estimate the following equations to compare the quantity purchased just before sales and just after sales:

$$\log(X^2_{H,ks}/X^1_{H,ks}) = c + \varepsilon_{ks}, \tag{4}$$

$$\log(X^2_{H,ks}/X^1_{H,ks}) = c + \varepsilon_{ks} + a_1 D^1_{week} + a_2 D^2_{week}, \tag{5}$$

$$\log(X^2_{H,ks}/X^1_{H,ks}) = c_j + \varepsilon_{ks}, \tag{6}$$

$$\log(X^2_{H,ks}/X^1_{H,ks}) = c_j + \alpha \log(T_{ks}) + \beta \log((P^1_{L,ks} + P^2_{L,ks})/P_{H,ks}) + \varepsilon_{ks}, \tag{7}$$

where j represents a product cateogry to which goods k belong. We control weekdays by introducing D^j_{week} $(j = 1, 2)$. It represents the days of week dummy, where $j = 1$ and 2 indicate those at the days just before and just after sales, respectively. Each D^j_{week} consists of 6 variables (e.g., one variable takes one when Monday, and another takes one when Tuesday).

We expect to obtain a negative c, suggesting that the quantity purchased just after sales is smaller than that just before sales, although the price is the same at P_H. Moreover, we control heterogeneity by introducing the fixed effect. Although this makes the interpretation of c no longer meaningful, we check whether α and β are negative and positive, respectively. In other words, we ask whether a longer duration of sales and a larger size of sales discounts lead to a bigger decrease in the quantity purchased just after sales.

Furthermore, we estimate the following equations to compare the quantity purchased during sales:

$$\log(X^2_{L,ks}/X^1_{L,ks}) = c + \varepsilon_{ks}, \tag{8}$$

$$\log(X^2_{L,ks}/X^1_{L,ks}) = c + \varepsilon_{ks} + a_1 D^1_{week} + a_2 D^2_{week}, \tag{9}$$

$$\log(X^2_{L,ks}/X^1_{L,ks}) = c_j + \varepsilon_{ks}, \tag{10}$$

$$\log(X^2_{L,ks}/X^1_{L,ks}) = c_j + \alpha \log(T_{ks}) + \beta \log(P^2_{L,ks}/P^2_{L,ks}) + \varepsilon_{ks}. \tag{11}$$

We expect to see negative c.

Figure XXX

Fact 3. Consumption tends to decrease as the inventory decreases.

Earlier studies on household inventory often assume that consumption is constant, but Hendel and Nevo (2006a) and Kano (2018) emphasize that consumption is state dependent. To confirm this, we employ the following exercise using the Shoku-map data. Particularly, we investigate whether consumption decreases until next purchase, as the inventory decreases. Suppose that household i uses goods k in date $t_l \in t_1, t_2, \cdots, t_{n_{ikt}}$, where t represents the purchase date. If she uses goods k twice in date t_{l_0},we record t_{l_0} twice. Thus, n_{ikt} represents the number of uses. We define the inventory in date $t^{'}$ $(t \leq t^{'} \leq t_{n_{ikt}})$ as $\lambda_{ikt'} = n_{ikt} - \sum_{l=1}^{n_{ikt}} 1$. Note that $\lambda_{ikt'}$ is an integer between 0 and n_{ikt}. For example, $\lambda_{ikt'} = n_{ikt}$ for $t \leq t^{'} \leq t_1$. We further define the sum of inventory Λ_{ijt} for household i in product category j at the beginning of date t as

$$\Lambda_{ijt} = \sum_{k \in j} \lambda_{ikt}. \tag{12}$$

Then, we estimate the following equation for the sample of $\Lambda_{ijt} > 0$:

$$y_{ijt} = c_i + d_j + \alpha \Lambda_{ijt} + \varepsilon_{ijt}, \tag{13}$$

where y_{ijt} is consumption. It takes one if goods in product cateogry j are used by household i in date t and zero otherwise. To save time, we estimate the model linearly without employing the logit or probit model. Alternatively, we define y_{ijt} as the sum of uses of goods in product cateogry j by household i in period t.

As Table 2 shows, the cofficient on the inventory α is significantly positive irrespective the choice of the dependent variable. Consumption is state dependent, decreasing as the inventory decreases. Endogeneity leads to a bias if ε_{ijt} is auto-correlated. In particular, if $\mathrm{Cor}(\varepsilon_{ijt}, \varepsilon_{ijt-1}) > 0$, then high consumption demand at $t-1$ decreases inventory Λ_{ijt} at t as well as increases consumption at t. This generates a downward bias in the estimate of α. Thus, this endogeneity, if present, would reinforce our result.

Moreover, our result is robustly confirmed in category-level estimation. We estimate the above equation for each cateogy allowing for variations in α for different cateogry j. Figure 6 shows that the distribution of t statistics for coefficient α_j. The vertical axis indicates the number of categories whose t statistics is larger than t^*, while the horizontal axis represents t^*. In 144 categories out of 157, t statistics for α_j is larger than 1.96 when the dependent variable is the consumption that takes one if used. It is 146 when the dependent variable is the number of uses.

3 Quasi Dynamic Model

3.1 Setup

In this section, we construct a simple partial-equilibrium model with the aim at explaining the stylized facts in the previous section, inferring inventory and consumption, and constructing the COLI. Product $k \in K_t$ is storable and, for simplicity, does not depreciate.

The novelty of our model is in the sepration of warehouse firms from households. Warehouse firms, who may be interpreted as a special member of households, have technology to hold inventory i_t^k, although there is a cost. They purchase storable goods from producers by x_t^k, hold inventory i_t^k, and sell them to households. Entry is free so that the expected firm value is zero. In this respect, there is no loss of consumer surplus. On the other hand, households in a narrow sense cannot hold inventory: their purchase always equals their consumption c_t^k. They purchase goods from warehouse firms and/or producers at the *effective* price r_t^k.

This framework enables us to solve the COLI in a conventional static manner. All we need to know comes down to two variables: c_t^k and r_t^k. We do not need to use complex dynamic COLI. Although the economy is admittedly fictional, many key properties continue to hold. For example, the three stylized facts can be explained not only by our quasi dynamic model but also by a dynamic model like the one used in Hendel and Nevo (2006a), although we do not show here. At the other extreme, a completely static model is problematic, because it is hard to explain the three stylized facts by neglecting storability.

We consider the optimization problems of the warehouse firms and the households. The price of storable goods, p_t^k, takes one of the two: high (regular) or low (sales) stochastically and exogenously.

Households

There are a unit mass of households. A representative household's cost minimization problem is expressed as

$$\min_{c_t^k} \left\{ \sum_{k \in K_t} r_t^k c_t^k + \lambda_t \left\{ U - \left[\sum_{k \in K_t} b_t^k \left(c_t^k \right)^{\frac{\sigma-1}{\sigma}} \right]^{\frac{\sigma}{\sigma-1}} \right\} \right\}, \tag{14}$$

where U is target utility, σ denotes the elasticity of substitutions, and $b_t^k = b^k + \varepsilon_t^k$ represents a taste or quality parameter for product k in period t with time-varying

flucuations of ε_t^k.[12] We extensively use the following relation with respect to the optimal quantity purchased:

$$c_t^k = \left(\frac{r_t^k / b_t^k}{r_{t'}^k / b_{t'}^k} \right)^{-\sigma} c_{t'}^k. \tag{15}$$

Warehouse Firms

There exist an infinite number of warehouse firms under free entry. Each warehouse firm maximizes the firm value:

$$V(i_{t-1}, p_t) = \mathrm{E}_t \left[\sum_{j=0}^{\infty} \beta^j \left\{ \sum_{k \in K_{t+j}} \left(r_{t+j}^k y_{t+j}^k - p_{t+j}^k x_{t+j}^k - C(i_{t+j}^k) \right) \right\} \right], \tag{16}$$

subject to the cost of inventory: $C(0) > 0$, $C(i)' > 0$, $C(i)'' \geq 0$, and the evolution of inventory:

$$i_t^k = i_{t-1}^k - y_t^k + x_t^k. \tag{17}$$

The firm sells goods k to the household by the amount of y_t^k with the price of r_t^k. Furthermore, the inventory as well as the purchase must be nonnegative:

$$x_t^k,\ i_t^k \geq 0. \tag{18}$$

The first-order conditions with respect to x_t^k and i_t^k are

$$0 = r_t^k - p_t^k + \psi_t^k, \tag{19}$$

$$C'(i_t^k) = \beta \mathrm{E}_t[r_{t+1}^k] - r_t^k + \mu_t^k, \tag{20}$$

where ψ_t^k and μ_t^k represent the Lagrange multipliers with respect to x_t^k and i_t^k, respectively. They are strictly positive when x_t^k and i_t^k are zero, respectively. They equal zero when x_t^k and i_t^k are positive.

The free entry condition leads to the nonpositive value for the entrant firm with zero inventory holding:

$$V(i_{t-1} = 0, p_t) \leq 0. \tag{21}$$

[12] There is no household heterogeneity in the model. Many previous papers, such as Boizot, Robin, and Visser (2001) and Hendel and Nevo (2004, 2006a), investigate the frequency of purchase conditional on past purchase behaviors. To obtain quantitatively plausible values for this, we need household heterogeneity, because otherwise the frequency of purchase would be either zero or one. In our study, we neglect household heterogeneity to focus on macroeconomic implications rather than individual behaviors.

Prices

The price of storable goods follows a Markov process. It takes either of the two, non-sale P_H or sale P_L ($P_H > P_L$) and

$$\begin{aligned} \text{Prob}(P_L|P_H) &= \overline{q} \\ \text{Prob}(P_L|P_L) &= \underline{q}. \end{aligned} \tag{22}$$

Market clearing

The goods market is cleared as

$$\int_0^{N_t} y_{t,j}^k dj + \int_0^{M_t} z_{t,j}^k dj = \int_0^1 c_{t,j}^k dj, \tag{23}$$

where $z_{t,j}^k$ represents the direct supply of storable goods k by producers to households. Household consumption equals household purchase from warehouse firms and producers. In the market, there are a unit mass of households, N_t warehouse firms, and M_t producers.

Note that warehouse firms we consider in the model are still a part of households, although we separate them to simplify our analysis. Thus, the quantity purchased that is recorded in the POS data, X_t, should equal the sum of the quantity purchased by warehouse firms $\int_0^{N_t} x_{t,j}^k dj$ and that purchased directly by households $\int_0^{N_t} z_{t,j}^k dj$. Clearly, this is not equal to aggregate consumption $\int_0^1 c_{t,j}^k dj$.

COLI

As we emphasize, the household's problem is static, so the COLI is the same as the conventional one. The cost minimization problem subject to constant utility leads to the optimal quantity purchased as

$$c_t^k = \left(r_t^k / b_t^k\right)^{-\sigma} \lambda_t^{\sigma} U. \tag{24}$$

The COLI, $\lambda_t = \mathcal{C}(r_t)$, is given by

$$\mathcal{C}(r_t) = \sum_{k \in K_t} r_t^k c_t^k = \left[\sum_{k \in K_t} \left(b_t^k\right)^{\sigma} \left(r_t^k\right)^{1-\sigma} U \right]^{1/(1-\sigma)}. \tag{25}$$

Although b_t^k is unobservable, equation (15) tells us that we need to know only two variables in every period: the effective price r_t^k and the share $r_t^k c_t^k$, to calculate the COLI

except for a constant term related to U.[13]

3.2 Equilibrium Properties

We discuss equilibrium properties by touching upon the aformentioned stylized facts. In the following discussions, we omit the superscipt of product k for simplicity. Denote the aggregate inventory at the end of period $t-1$, by $I_{t-1} \equiv \int_0^{N_t} i_{t-1,j} dj$.

The first lemma states the property of the effective price between warehouse firms and households, r_t. Unless warehouse firms hold excessive inventory due to un expected shock, r_t equals P_L when p_t equals P_L (sales). Households purchase goods directly from producers at $r_t = P_L$. When $p_t = P_H$ (regular), households may purchase goods from warehouse firms at a price below P_H. The price r_t is lower as the inventory I_{t-1} is larger.

Lemma 1 *The effective price r_t satisfies $0 < r_t \leq P_H$. Suppose no big unexpected shock to b_t. Then, r_t satisfies $P_L \leq r_t \leq P_H$ and, when $p_t = P_L$, $r_t = P_L$. Furthermore, $r_t = r(I_{t-1}, p_t, b_t)$ is non-decreasing in p_t and b_t and non-increasing in I_{t-1}.*

The proofs of this lemma and those that follow are all provided in the Appendix. The next lemma states the stockpiling behavior of warehouse firms. Only when $p_t = P_L$, they purchase goods and hold inventory with the aim at selling them at a higher price after sales end.

Lemma 2 *Suppose no big unexpected shock to b_t. If $p_t = P_H$, warehouse firms do not purchase, that is, $x_t = 0$. If $p_t = P_L$, warehouse firms purchase and hold inventory, which is independent of i_{t-1}, I_{t-1}, and b_t.*

The left-side panel of Figure 7 illustrates the pattern of price and quantity changes in a sales event, particularly, when the inventory is held just for one period. The top and bottom panels show prices and quantities, respectively. Sales occur at $t = 2$ and 3, while price is high in other periods. The bottom panel shows that, at $t = 2$, the quantity purchased, X_t, represented by the filled circle becomes the largest. This is not just because the price is lower but also because warehouse firms stockpile. Thus, X_t is higher than household consumption, c_t, shown in the circle in red by the size of inventory.

[13] Considering that warehouse firms are a part of households, it would be clearly desirable to embed the COLI for warehouse firms' perspective. However, this effect is considered to be small because the free entry condition makes the value of warehouse firms zero.

At $t = 3$, X_t coincides with c_t because there is no additonal need for stockpiling. Then, at $t = 4$ when sales end, warehoues firms sell their inventory to households. The effective price traded r_t lies between P_L and P_H. Since consumption is decreasing with the effective price, its level at $t = 4$ is lower than that at $t = 2$ and 3 but higher than that at $t = 1$ and 5. At the beginning of $t = 5$, warehouse firms hold no inventory, so households purchase goods at the price of P_H and consume less.

Stylized Fact 2

The following lemma explains Stylized fact 2.

Lemma 3 *Suppose no shock: $b_t = b$. Then, the quantity purchased by warehouse firms and households just before sales is greater than or equal to that just after sales.*

The quantity purchased by warehouse firms and households at the first day of sales is greater than or equal to that at the final day of sales.

One remark is worth making. In the model, prices are stochastic. Warehouse firms and households cannot accurately predict future prices. This is why warehouse firms hold inventory at the first day of sales. Instead, suppose that the duration of sales is known *ex ante*. Then, there is no incentive for warehouse firms to stockpile except for the last day of sales. The quantity purchased by warehouse firms and households at the first day of sales should become smaller than or equal to that at the final day of sales.

Stylized Fact 3

Consumption c_t is clearly state dependent, particularly, on effective price r_t.

Lemma 4 *Suppose no shock: $b_t = b$. Then, c_t is decreasing with r_t. After sales end, r_t and c_t are non-decreasing and non-increasing over time, respectively, until the next sales begin.*

Stylized Fact 1

Our model succeeds in explaining the chain drift.

Lemma 5 *Consider one sales event for product k such that $p_t = p_{t+T+1} = P_H$ and $p_{t+j} = P_L$ for $j = 1, \cdots, T$ ($T \geq 1$). Suppose that the effect of product k is so small that the price and quantity of other goods are unchanged; $\sum_{k' \in K_0 \cap K_t} p_t^{k'} x_t^{k'} = 1$; the price*

before and after sales is P_H for a sufficiently long duration (i.e., $p_{t-j} = p_{t+T+1+j} = P_H$ for $j = 0, 1, \cdots, T_H$, where T_H is sufficiently large compared with m); and $b_t^k = b^k$.

If $\sigma > (<)1$ and $m \geq 1$, the change in the price index from t to $t+T+1+T_H$ satisfies $\pi^{COLI} = 0$; $\pi^L > (<)0$; $\pi_t^T < 0$; $\pi^{T*} < (>)0$. If $\sigma = 1$ or $m = 0$, all the above change in the price index is zero except for $\pi_t^T < 0$.

Furthermore, if $\sigma > 1$, $\pi^P < \pi^T < 0 < \pi^L$. If $\sigma > 1$ and $m = 1$, $\pi^T < \pi^{T*} < 0$.

In the above, m represents the degree of stockpiling, which indicates how long inventory remains at the hand of warehouse firms after sales end. In the next section, we will specify how m is determined. In the Appendix, we discuss how equilibrium is determined in a special case of $m = 1$, that is, when inventory is cleared in just one period after sales.

Superlative Indexes

It is well known that the Törnqvist index is a good approximation of the COLI up to the second order (Diewert (1976)). While the purchase-based Törnqvist index is clearly bad in the presence of storable goods, the above lemma also suggests that the use of consumption cannot eliminate the chain drift.

The reason is systematic asymmetry due to the non-negative constraint of inventory and purchase. As Figure 7 shows, the movement of the effective price is asymmetric between price decrease and price increase. On the one hand, a price decrease causes an instantenous decrease in the effective price. On the other hand, the response of the effective price to a price increase is gradual. Such asymmetry is not a coincidence, but a natural outcome of the non-negative constraint of inventory and purchase. Thus, the third-order error comes to matter, although the Törnqvist index is a good approximation of the COLI up to the second order.

To resolve the chain drift and derive a better approximation of the COLI, we need a superlative index with the consideration of the elasticity of substitution σ. This is the order r superlative index, where we define P_r as

$$P_r(r_0, r_1, c_0, c_1) = \frac{\left\{\sum_{k\in K} s_0^k \left(\frac{r_1^k}{r_0^k}\right)^{(1-\sigma)}\right\}^{1/\{2(1-\sigma)\}}}{\left\{\sum_{k\in K} s_1^k \left(\frac{r_0^k}{r_1^k}\right)^{(1-\sigma)}\right\}^{1/\{2(1-\sigma)\}}}, \tag{26}$$

where s_t^k represents the consumption share of product k at t.

The following lemma shows that P_r serves as the COLI if the utility is expressed as

$$\mathcal{C}(r_t) = \left[\sum_{i \in K} \sum_{k \in K} \alpha^{ik} \left(r_t^i\right)^{(1-\sigma)} \left(r_t^k\right)^{(1-\sigma)} \right]^{1/\{2(1-\sigma)\}} \tag{27}$$

where $\alpha^{ik} = \alpha^{ki}$. This is a more generalized form of utility than that given by equation (14).[14]

Lemma 6 *Given the unit cost function of (27), P_r equals $\mathcal{C}(r_1)/\mathcal{C}(r_0)$.*

4 Inference of Consumption and the Calculation of the COLI

4.1 Aim

Using the model, we have succeeded in explaining the stylized facts. However, many important variables such as consumption, inventory, and the effective price remain unobservable. From a practical perspective, consumption and the effective price are essential to construct the COLI. For macroeconomists, it is of great interest as to how stockpiling behavior has changed over time and what its determinants are. Thus, in this section, we aim to infer these variables using a simple, tractable approach. Furthermore, it is important to note that the estimate of the elasticity of substitution is biased unless we properly take account of stockpiling. Even if consumption and effective prices have a clear negative relation characterized by σ as is shown in equation (15), it does not mean that observable purchase and prices have the same clear relation. As Hendel and Nevo (2004) argue, price reductions influence the quantity purchased via not only a consumption effect (consumption is price sensitive) but also a stockpiling effect (i.e., consumers stockpile for future consumption). Our approach enables us to estimate σ.

4.2 Our Approach

Hereafter, we suppose no aggregate demand (taste or quality) change for each product: $b_t = b$. Then, the key variable is the effective price at $p_t = P_H$ after sales end, that is, $r_H(I_{t-1}) \equiv r(I_{t-1}, P_H)$, where $r_H(0) = P_H$. The effective price at $p_t = P_L$ equals P_L.

[14] The Lloyd–Moulton Index also serves as the COLI. See ILO et al. (2004).

Knowing this variable enables us to obtain consumption c_t after sales using equation (15).

When $p_t = P_H$, warehouse firms optimize as

$$C'(i_H; I_{t-1}) = \beta \{(1-\overline{q}) r_H(I_t) + \overline{q} P_L\} - r_H(I_{t-1}) + \mu_t \tag{28}$$

from equation (20). This equation shows that warehouse firms strike a balance between the benefit from a future effective-price increase (right-hand side) and the cost from holding inventory (the left-hand side). Note that μ_t is the Lagrange multiplier associated with i_t and equals zero if $i_t > 0$.

In a special case, we notice that $r_H(I_t) - r_H(I_{t-1})$ becomes positive constant. In other words, the expectation of a *linear* effective-price increase prevents warehouse firms from selling all of their inventory instantaneously and selling none. Warehouse firms gradually sell off their inventory to households. See Appendix for the condition this linearity holds.

In the following analysis, we assume the linearity. Admittedly, this is a very restrictive assumption, but at the same time, can be interpreted as an intermediate of the following two scenarios. One is when $C'' = 0$. In this scenario, $r_H(I_t) - r_H(I_{t-1})$ is increasing with t. Put differently, if inventory cost is not convex, warehouse firms require a greater effective-price increase as time goes by, because they discount the future ($\beta < 1$) and expect the next sales to come again ($\overline{q} > 0$). On the other hand, suppose $\beta = 1$ and $\overline{q} = 0$. Then, $r_H(I_t) - r_H(I_{t-1})$ is decreasing with t. The cost of holding inventory decreases as their inventory decreases because of $C'' > 0$, which makes warehouse firms require a smaller effective-price increase as time goes by.

This linearity assumption greatly simplifies our analysis. Given the path of $r_H(I_{t-1})$, we compute consumption c_t as $(r_H(I_{t-1})/P_L)^{-\sigma} c_L^*$, where c_L^* represents consumption during sales, which we specify below. Furthermore, we can calculate the degree of stockpiling, m. More precisely, we define m to denote the duration of inventory after sales end. In continuous time, we can derive the following equation:

$$m_{cont} = \frac{P_H - P_L}{P_L} \frac{\sigma - 1}{1 - (P_H/P_L)^{-\sigma+1}} \frac{I_L}{c_L^*}. \tag{29}$$

See Appendix for the proof. Simply put, this can be derived because accumulated consumption for m_{cont} periods equals the initial inventory oustandings just after sales end, I_L, and the effective price linearly increases from $r_H(t+T) = P_L$ to $r_H(t+T+m_{cont}) = P_H$, when sales end at $t+T$. The right-side panel of Figure 7 illustrates the pattern of price and quantity changes in the case of $m = 5$. The effective price depected with the unfilled

circle in red in the top panel increases at a constant rate from $t = 3$ to 8. The bottom panel shows that, while the quantity purchased falls to zero from $t = 4$ to 7, consumption is not zero, decreasing gradually.

In the Appendix, we provide detailed explanations about how we calculate $r_H(I_{t-1})$ and m as well as c_L^* and I_L for each sales event of product k and shop s. Briefly put, c_L^* is the lower value of the quantity purchased (observable in the POS data) during sales. Note that, according to the model, consumption during sales is equal to purchase except for the first day of sales, since there is no need for further stockpiling. The inventory I_L equals the cumulative amount of the purchase during sales minus Tc_L^*.

To calculate σ for each 3-digit product category, for each sales event of product k and shop s, we collect the record of the log ratio of quantity consumed at sales to that at regular divided by the log ratio of price at sales to that at regular. Equation (15) suggests that this equals σ on average.[15] We calculate the unweighted average of Γ across sales events, products and shops for each 3-digit product category, which we define as σ. This approach is clearly different from a naive approach that directly uses observable data on price (p_t) and quantity (X_t) to calculate $-\log(X_t)/\log(p_t)$ or its time difference.

4.3 Simulation Results

In this subsection, we illustrate equilibrium by employing numerial simulation. We simulate the paths of the effective price, consumption, and purchase as well as the chain drift by generating randomazied price paths. We then examine if we can reproduce the pattern of quantity purchased and the chain drift as was observed in the POS data and predicted by the model. Moreover, we study how the degree of the chain drift changes depending on the parameter values associated with sales and stockpiling.

The model is based on that introduced in the previous section. Instead of specifying the form of the cost of inventory $C(i)$, we specify m and calculate other variables such as inventory, consumption, and the effective price so as to be consistent with this specification. We use the following parameter values. There are $K = 100$ storable goods with identical characteristics. Regarding sales, their probability is governed by $\overline{q} = 0.03$ and $\underline{q} = 0.50$ and their size is given by $P_L/P_H = 0.9$. These values are chosen to be roughly consistent with our POS data. The degree of stockpiling m and the elasticity

[15] It is natural to assume that prices are exogenous for households, in particular, in such a short time horizon. Thus, a bias coming from endogeneity is not a big problem.

of substitution σ are not directly observable in our data, but the above approch enables us to infer them. We set them at 5 and 4, respectively. We generate randomazied price paths for periods $T = 405$ days and discard the first and last 20 samples to calculate changes in the price indexes in 365 days. We repeat this for $N = 100$ times.

The left-hand panel of Figure 8 visulaizes a typical path of price and quantity, which we draw by sampling certain $t's \in T$, $k \in K$, and $n \in N$. This confirms Stylized Fact 2: the changes in the quantity purchased in a sales event. The quantity purchased falls to zero for m periods after sales end, which is smaller than that just before sales. When sales last two days, the quantity purchased is greater in the first day than in the second day. For m periods after sales, the effective price r_t increases at a constant rate from P_L to P_H, while consumption decreases nonlinearly because $\sigma > 1$.

Next, we calculate the price indexes based on the COLI, chained quadratic mean of r, chained consumption-based Törnqvist, chained purchase-based Törnqvist, chained purchase-based Laspeyres, and chained purchase-based Paasche. For each $n \in N$, we calculate the price levels after 365 days by setting initial price levels at one. We then take the mean and the standard deviation of them for n.

Table 3 shows the simulation results. The COLI and chained quadratic mean of r are exempt from the chain drift, going back to the same level asymptotically. Chained consumption-based Törnqvist has a downward chain drift by 1% annually. Purchase-based price indexes have far greater chain drifts: Törnqvist decreases by about 84% annually, Laspeyres increases by about 27, and Paasche decreases to almost 10^{-3}. The size of the chain drift for purchase-based Törnqvist is comparable to the actual size shown in Figure 1. Furthermore, this table shows that the size of the chain drifts increases, as the degree of stockpiling, m, increases, the size of sales discount increases (P_L/P_H decreases), or the probability of sales given non-sales, $\overline{q}$, increases. By contrast, it is intereting to point out that an increase in the probability of sales given sales, $\underline{q}$, decreases the size of the chain drifts for purchase-based Laspeyres and Paasche, albeit not purchase-based Törnqvist. This is considered to arise because the expectations of longer continuation of sales decreases an incentive of stockpiling.

4.4 Applying Our Approach to Japan's Retailer Scanner Data

We apply the approach explained in Section 4.2 to Japan's POS scanner data. The right-hand panel of Figure 8 shows the actual paths of the price and the quantity purchased

for a particular instant cup noodle at a shop. While these paths are observable, the paths of consumption and the effective price are not. Our approach enables us to infer them, as shown in the figure.

Figure 9 shows the histogram of the degree of stockpiling m at the 3-digit product category level. It distributes from 1.0 to 5.6 with the mode being around 2. The three top categories are instant cup noodles, diluted beverages, and frozen meals in the descending order. These products are indeed storable for a relatively long duration. However, the durability of products does not neccesarily imply a high degree of stockpiling. In fact, the three bottom categories are razors, chilled condiments, and cosmetic accessories. Except for chilled condiments, they are highly durable.

The degree of stockpiling is highly heterogeneous within a product category. Taking a sample of m from each product, shop, and sales event, we draw the cumulative distribution for the degree of stockpiling ($m < m^*$) for the cateogories of instant cup noodles and tofu products. Figure 10 shows that, while around half of the sample exhibits the lowest value of m (i.e. $m = 1$), instant cup noodles have a thicker right tail than tofu products. Because tofu products expire sooner, they are stockpiled less at sales than instant cup noodles.

Sigma estimate

4.5 External Validations

4.5.1 Comparison of the Inflation Rates: Data and Simulation

To validate our approach, we examine whether the size of simulated chain drift is comparable to that of actual one. As for data, we calculate the daily average of the inflation rate from January 1989 to December 2011 based on the purchase-weighted Törnqvist index for each 3-digit product category j (denote π_j). About 12 months of the data are omitted because there seems a bias. At the same time, we record the average of the following variables for each 3-digit product category: the degree of stockpiling m_j, the probability of sales given non-sales $\overline{q_j}$, the probability of sales given sales $\underline{q_j}$, and the size of sales discount $(P_L/P_H)_j$. Then, using the values, we simulate the model and calculate the average size of the chain drift as in Section 4.3. The elasticity of substitution σ is set at 5 for all the product categories, which hardly influences the size of the chain drift as shown in Table 3.

Figure 12 shows that our approach does a reasonally good job in explaining the chain

drift. In the figure, each circle indicates a product category. It shows that a product category that has a large chain drift in the data tends to exhibit a large chain drift in the simulation as well. The circles tend to lie below the 45 degree line, suggesting that the size of the actual chain drift is larger than that of the simulated one. However, the quantitative difference is not big except for two categories, frozon staple foods and frozon meals, that entail large chain drift in the simulation.

Next, we regress the average inflation rate π_j based on data by $\log(m_j)$, $\overline{q_j}$, $\underline{q_j}$, and $\log(P_L/P_H)_j$. Table 4 shows the estimation result. The coefficients on $\log(m_j)$ and $\overline{q_j}$ are negative and significant, that on $\log(P_L/P_H)_j$ is positive and significant, and that on $\underline{q_j}$ is insignificant. This result is all consistent with the simulation result shown in Table 3: high m, high $\overline{q}$, and low P_L/P_H induces greater deflation, while the effect of $\underline{q}$ is negligible.

4.5.2 Comparison with Shokumap Data

The next way to validate our approach is to compare the degree of stockpiling (m) obtained from the POS data with the related variables in the Shokumap data. As for the latter, we look at the duration of inventory defined by the difference between the date of puchase (t_p) and the date of the last consumption (t_l). Using the cross-sectional dipersions of m at the 3-digit product category level, we examine if there is any significant correlation. The left-side panel of Figure 13 shows that there is no significant correlation between $\log(m)$ in the POS data and $\log(t_l - t_p + 1)$ in the Shokumap data. This suggests that m does not necessarily represent storability.

However, a significant correlation arises when we look at a change in $\log(m)$ in March 2014. On April 1st, 2014, the consumption tax rate was raised from 5% to 8%. This was a preannounced event, which prompted households to stockpile before the tax increase. In fact, the right-side panel of Figure 13 shows that the change in $\log(m)$ in March 2014 from that in the same month in the previous year is mostly positive. Furthermore, a correlation coefficient between this variable and $\log(t_l - t_p + 1)$ is as high as +0.41. This evidence is in favor of our interpretation of m: it captures a degree of stockpiling. The figure implies that the degree of stockpling increases in response to anticipated price increase and that goods with a longer inventory duration tend to be stockpiled more.

If m captures stockpiling, it should embed not only storability but also the quantity purchased. For the reason, the average amount of purhcase in the Shokumap data, $\log(q)$,

may correlate with $\log(m)$ in the POS data. Figure 14 shows that this is indeed the case. Positive correlations are observed between $\log(q)$ and $\log(m)$ both for the average of the entire observation periods and in March 2014. The correlation coefficient is greater for the latter period (+0.45) than the former (+0.31).

4.6 Price Indexes

Using the POS data, we calculate the time-series of the price level based on the following three types of definitions. The first one is the Törnqvist index based on the purchase weight. This is conventional and we previously showed in Figure 1. Second, the price index is the Törnqvist index based on the consumption weight. Here, we use the effective price as well as consumption to calculate the price index. The third one is the order r superlative index, where we use the estimated elasticity of substitution σ at a category level.

Figure 15 shows the time-series paths of the three price indexes, where the innitial price level is normalized as one. The huge downward bias in the conventional Törnqvist index is mitigated greatly by consumption-based indexes. However, the downward movement is not completely eliminated. Whereas the order r superlative index yields a slightly higher price change, both consumption-based indexes exhibit deflation by approximately 10% annually.

Such deflation calls for futher improvements of our approach, although this may not necessarily mean a bias. For example, more careful treatments are needed in, for example, quality adjustment and price imputations. Moreover, recall that our data frequency is daily. Thus, even a tiny size of chain drift becomes large, when it is is accumulated from daily to a longer time horizon. For example, a price chage by -10% annually amounts to a price chage by $((1-0.1)^{1/365}-1)\cdot 100 = -0.0288\%$ daily only.

4.7 Developments in Households' Stockpiling Behavior

4.7.1 Changes in Stockpiling Behavior

The degree of stockpiling is not only heterogenous but also time varying. Figure 16 shows XXX

It should be noted that m can change by a change in not only households' intrinsic behavior but also prices that are exogenous to households. To see this formally, we use

the previous model. When $p_t = P_L$, warehouse firms optimize as

$$C'(i_L; I_{t-1}) = \beta \left\{ (1 - \underline{q}) r_H(I_t) + \underline{q} P_L \right\} - P_L + \mu_t \tag{30}$$

from equation (20). This equation tells how much warehouse firms should hold inventory at sales. If sales end in period t, the effective price increases in period $t+1$, which provides them a profit if they hold inventory. However, it incurs a cost of holding inventory, as the left-hand side of the equation shows. Suppose $C'(i_t; I_{t-1}) = C > 0$ and the linear effective-price increase in r_H. Then, we have

$$C = \beta(1 - \underline{q}) \left\{ \frac{1}{m} \frac{P_H - P_L}{P_L} + 1 \right\} P_L - (1 - \beta \underline{q}) P_L,$$

and, in turn,

$$m = \beta(1 - \underline{q}) \frac{P_H - P_L}{P_L} \left(1 - \beta + \frac{C}{P_L} \right)^{-1}. \tag{31}$$

This equation suggests that the degree of stockpiling m depends on the probability of sales at t conditional on sales at $t-1$, negatively, the size of sales discounts positively, and the cost of inventory negatively. In other words, more stockpiling occurs, as sales are expected to end sooner and the size of sales discounts is greater. More generally, r_H is considered to depend on $\overline{q}$ as well.

Thus, we regress the following equation for $\log(m_{jt})$:

$$\log(m_{jt}) = c_j + d_t + AX_{jt} + \varepsilon_{jt}, \tag{32}$$

where X_{jt}= $\{\log(1 - \overline{q_{jt}}), \log\left(1 - \underline{q_{jt}}\right), \log\left((P_H - P_L)/P_H\right)_{jt}\}$ and j and t represent a 3-digit product category and a month, respectively. Time fixed effects d_t are included to capture the aggregate, demand-side, time-varying component of m_{jt}.

Table 6 shows the estimation results. The coefficient on $\log(1 - \overline{q})$ is insignificant when time fixed effects are included, which is consistent with equation (31). Also consistent is the coefficient on $\log((P_H - P_L)/P_H)$, which is positive and significant. However, the coefficient on $\log\left(1 - \underline{q}\right)$ is negative, which is opposite to equation (31). One possible reason is the endogeneity of $\underline{q}$. It should be noted that this regression does not necessarily help investigate causality because the cost of inventory in equation (31) is not controled. For a product category with low inventory costs, m tends to be high. If firms hold longer sales (high $\underline{q}$) for these products, it is expected to observe a negative coefficient on $\log\left(1 - \underline{q}\right)$ rather than positive. Another reason is that, in equation (29),

m is proportional to I_L, which tends to increase as the duration of sales T increases. Owing to this construction, our measure of m tends to increase as $\underline{q}$ increases.

In Figure 16, we plot time fixed effects d_t, a change in the degree of stockpiling with the effects of price changes being excluded.

4.7.2 Dependence on Business cycles

Does the stockpiling behavior of households have implications for macroeconomics? To address this question, we examine whether the degree of stockpiling m depends on business cycles. We estimate the following equation:

$$\log(m_{jt}) = c_j + Au_t + \mu_{jt}, \tag{33}$$

where u_t represents the unemployment rate and/or hours worked. Coefficients are considered to be both positive and negative. On the one hand, since the opportunity cost of shopping relative to working increases during an economic boom, greater stockpiling may occur. If so, a coefficient on the unemployment rate is negative. Furthermore, a less stringent financial constraint may induce greater stockpiling in booms. Conversely, households go shopping more frequently and purchase less volatile amounts in recessions. However, the opposite sign is possible. Households may become more price sensitive in a recession than in a boom. If so, households stockpile more in recessions, and a coefficient on the unemployment rate becomes positive.

Table 7 shows that the degree of stockpiling depends on business cycles. We find that the coefficients on the unemployment and hours worked are negative and positive, respectively, and they are both significant. Thus, the story based on the opportunity cost of shopping or the financial constraint seems to fit. Moreover, Columns (4) and (5) show that the coefficients remain significant even when we control pricing variables by including $\log(1-\overline{q_{jt}})$, $\log\left(1-\underline{q_{jt}}\right)$, and $\log\left((P_H-P_L)/P_H\right)_{jt}\}$. Thus, a degree of stockpiling depends on business cycles not just through changes in pricing by reatilers but also through changes in households' behavior on their own.

One may wonder how quantitatively important the business-cycle dependence of stockpiling is for macroeconomy. In Japan, it is considered to be truly important. Revenue from temporary sales is 30% of total revenue (Sudo et al. (2018)).

We calculate crude estimates on changes in the quantity purchased and the unit price by the change in m induced by business cycles. To this end, first, we use the above

estimated equation to evalulate the change in m in response to 1σ deviations in both the unemployment rate and hours worked. In a boom, a change in $\log(m)$ is estimated as $-0.0382 \cdot (-1.098) + 2.8795 \cdot 0.0387$. Second, using the approach explained in Section 4.3, we simulate the time-series path of the quantities purchased and the prices. We use the parameter values, such as $\overline{q}$ and $\underline{q}$, fixed at their cross-sectional and time-series means, while we use two different values for m. Third, we calculate the mean of the amount of sales (price times the quantity purchased), the quantity purchased, and the unit price (the amount of sales divided by the quantities purchased) for 365 days, 100 products, and 100 iterations.

Simulation results reveal that an endogenous increase in m from 2.2 to 2.5 increases the amount of sales by 0.8, increases the quantity purchased by 1.0%, and decreases the unit price by 0.1%. When the economy is in a good shape, households stockpile more at sales, which increases the average salesa and the quantity and decreases the unit price. Quantitatively, this size should be never negligible, although our simulation is calibrated only based on the products sold in supermarkets.

We estimate the equation corresponding to Column (5) of Table 7 but now for each 3-digit product category by allowing for variations in A for different category j. Figure 17 shows the distribution of t statistics. The vertical axis indicates the number of categories whose t statistics is smaller (larger) than t^*, while the horizontal axis represents t^*. The right panel shows that, in many categories (115 out of 145), t statistics for the coefficient on hours worked is larger than 1.96. For the coefficient on the unemployment rate, results are weaker: t statistics is lower than -1.96 only in 72 categories. In 16 categories, t statistics is higher than 1.96, showing a cyclicality in the opposite direction.

5 Conclusion

In this study, we documented

Future tasks include the more careful considerations about heterogeneity in terms of products as well as households. Improvements in the chain drift

References

[1] Abe, Naohito and Akiyuki Tonogi (2010) "Micro and Macro Price Dynamics in Daily Data," *Journal of Monetary Economics*, 57(6), 716–728.

[2] Blattberg, Robert, and Scott Neslin (1989) "Sales Promotion: The Long and the Short of It," *Marketing Letters*, 1(1), 81–97.

[3] Boizot, Christine, Jean-Marc Robin and Michael Visser (2001) "The Demand for Food Products. An Analysis of Interpurchase Times and Purchased Quantities," *Economic Journal*, 111, 391–419.

[4] Cashin, David, and Takashi Unayama (2016) "Measuring Intertemporal Substitution in Consumption: Evidence from a VAT Increase in Japan," *Review of Economics and Statistics*, 98(2), 285–297.

[5] Chevalier, Judith and Anil Kashyap (2019) "Best Prices: Price Discrimination and Consumer Substitution," *American Economic Journal: Economic Policy*, 11(1), 126–159.

[6] Diewert, W. Erwin (1976) "Exact and Superlative Index Numbers," *Journal of Econometrics*, 4(2), 115–145.

[7] Erdem, Tulin, Susumu Imai, and Michael Keane (2003) "A Model of Consumer Brand and Quantity Choice Dynamics under Price Uncertainty," *Quantitative Marketing and Economics*, 1(1), 5–64.

[8] Feenstra, Robert, and Matthew Shapiro (2003) "High-Frequency Substitution and the Measurement of Price Indexes," in *Scanner Data and Price Indexes*, Robert Feenstra and Matthew Shapiro (eds.), University of Chicago and NBER.

[9] Frisch, Ragnar (1936) "Annual Survey of General Economic Theory: The Problem of Index Numbers," *Econometrica*, 4(1), 1–38.

[10] Gowrisankaran, Gautam, and Marc Rysman (2012) "Dynamics of Consumer Demand for New Durable Goods," *Journal of Political Economy*, 120, 1173–1219.

[11] Griffith, Rachel, Aviv Nevo, Ephraim Leibtag, and Andrew Leicester (2009) "Consumer Shopping Behavior: How Much Do Consumers Save?" *Journal of Economic Perspectives*, 23, 99–120.

[12] de Haan, Jan and van der Grient (2011) "Eliminating Chain Drift in Price Indexes Based on Scanner Data," *Journal of Econometrics*, 161, 36–46.

[13] Hendel, Igal, Alessandro Lizzeri, and Nikita Roketskiy (2014) "Nonlinear Pricing of Storable Goods," *American Economic Journal: Microeconomics*, 6(3), 1–34.

[14] Hendel, Igal, and Aviv Nevo (2003) "The Post-Promotion Dip Puzzle: What do the Data Have to Say?" *Quantitative Marketing and Economics*, 1, 409–424.

[15] Hendel, Igal, and Aviv Nevo (2004) "Intertemporal Substitution and Storable Products" *Journal of the European Economic Association*, 2(2/3), 536–547.

[16] Hendel, Igal, and Aviv Nevo (2006a) "Sales and Consumer Inventory." *RAND Journal of Economics*, 37, 543–561.

[17] Hendel, Igal, and Aviv Nevo (2006b) "Measuring the Implications of Sales and Consumer Inventory Behavior," *Econometrica*, 74, 1637–1673.

[18] Hong, Pilky, Preston McAfee, and Ashish Nayyar (2002) "Equilibrium Price Dispersion with Consumer Inventories," *Journal of Economic Theory*, 105, 503–517.

[19] ILO, IMF, OECD, Eurostat, United Nations, and World Bank (2004) "Consumer Price Index Manual: Theory and Practice," ILO Publications, Geneva.

[20] Ivancic, Lorraine, W. Erwin Diewert, and Kevin J. Fox (2011) "Scanner Data, Time Aggregation and the Construction of Price Indexes," *Journal of Econometrics*, 161, 24–35.

[21] Kano, Kazuko (2018) "Consumer Inventory and Demand for Storable Goods: New Evidence from a Consumer Survey," *Japanese Economic Review*, 69(3), 284–305.

[22] Konus A. A. (1924) "The Problem of the True Index of the Cost of Living," in *Economic Bulletin of the Institute of Economic Conjuncture* (in Russian), 9–10, 64–71; published in English (1939) in *Econometrica*, 7, 10–29.

[23] Lloyd, P. J. (1975) "Substitution Effects and Biases in Nontrue Price Indices," *American Economic Review*, 65(3), 301–313.

[24] Nakamura, Emi, and Jon Steinsson. (2010) "More Facts about Prices." Columbia University. http://www.columbia.edu/en2198/papers/fivefactssupplement.pdf. Accessed February 28, 2017.

[25] Neslin, Scott, and Schneider Stone (1996) "Consumer Inventory Sensitivity and the Postpromotion Dip," *Marketing Letters*, 7, 77–94.

[26] Osborne, Matthew (2018) "Approximating the Cost-of-Living Index for a Storable Good," *American Economic Journal: Microeconomics*, 10(2), 286–314.

[27] Reinsdorf, Marshall B. (1999) "Using Scanner Data to Construct CPI Basic Component Indexes," *Journal of Business and Economic Statistics*, 17(2), 152–160.

[28] Reis, Ricardo (2009) "A Dynamic Measure of Inflation," Working Paper.

[29] Salop, Steven C. and Joseph E. Stiglitz (1982) "The Theory of Sales: A Simple Model of Equilibrium Price Dispersion with Identical Agents," *American Economic Review*, 72(5), 1121–1130.

[30] Sudo, Nao, Kozo Ueda, and Kota Watanabe (2014) "Micro Price Dynamics under Japan's Lost Decades," *Asian Economic Policy Review*, 9(1), 44–64.

[31] Sudo, Nao, Kozo Ueda, Kota Watanabe, and Tsutomu 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.

[32] Wang, Emily Yucai (2013) "The Effect of More Accurate Cost-of-Living Indexes on Welfare Recipients," Working Paper.

A Notes on Retailer Scanner Data

A.1 Aggregation of Variables

We aggregate variables of interest over days, products, and shops in the following way.

1. To construct the price indexes and the COLIs, we use the formula explained in the main text for products in each 3-digit product category. Note that we take the same product differently when it is sold in a different shop, because the timing of sales and the degree of stockpiling differ across shops. We then aggregate the price indexes and the COLIs at the 3-digit product category level using the sales weight in the month that includes the day.

2. Otherwise, we take the logarithm of a variable (unless it is a rate of change or ratio) and then aggregate the values over products and shops, assigning equal weights. The variables include $\log(m)$, $\log(1-\overline{q})$, $\log(1-\underline{q})$, and $\log((P_H-P_L)/P_H)$.

B Model Details

Proof of Lemmas 1 and 2

Define the Lagrangian as

$$\mathcal{L}(i_{t-1},p_t)=\max_{x_{t+j},i_{t+j}}\mathrm{E}_t\left[\sum_{j=0}^{\infty}\beta^j\left\{\begin{array}{c}(r_{t+j}(i_{t+j-1}-i_{t+j}+x_{t+j})-p_{t+j}x_{t+j}-C(i_{t+j}))\\ +\psi_{t+j}x_{t+j}+\mu_{t+j}i_{t+j}\end{array}\right\}\right]. \tag{34}$$

The first-order conditions with respect to x_t and i_t are

$$0=r_t-p_t+\psi_t, \tag{35}$$

$$C'(i_t)=\beta\mathrm{E}_t[r_{t+1}]-r_t+\mu_t. \tag{36}$$

The price to be purchased by the household, r_t, satisfies $0<r_t\leq P_H$. This is because, if r_t is strictly higher than P_H, the household purchases storable good k from a goods k producer directly. Thus, P_H serves as the price upper limit.

Furthermore, r_t generally lies in the range of $[P_L,P_H]$. If $\mathrm{E}_t\left[r_{t+1}\right]$ is strictly lower than P_L, warehouse firms suffer negative profits, because they purchase goods at P_L or

P_H. Anticipation of this possibility leads to no firm entry. However, a big *unexpected* shock may induce warehouse firms to do fire sales to save inventory costs. Without such a big unexpected shock, when $p_t = P_L$, we have $r_t = r(I_{t-1}, P_L, b_t) = P_L$ because the household can purchase storable goods from a goods producer directly.

Thus, our interest is particularly in the price at P_H, that is, $r(I_{t-1}, P_H, b_t)$. It is obvious to see $r(I_{t-1}, P_H, b_t) \geq r(I_{t-1}, P_L, b_t) = P_L$. When $p_t = P_L$ and P_H, equation (36) becomes

$$C'(i_L; I_{t-1}) = \beta \left\{(1 - \underline{q}) r(I_t, P_H, b_t) + \underline{q} P_L\right\} - P_L + \mu_t, \tag{37}$$

$$C'(i_H; I_{t-1}) = \beta \left\{(1 - \overline{q}) r(I_t, P_H, b_t) + \overline{q} P_L\right\} - r(I_{t-1}, P_H, b_t) + \mu_t, \tag{38}$$

respectively. According to equation (37), at $p_t = P_L$, inventory i_L is increasing in $r(I_t, P_H, b)$ and independent of i_{t-1}, I_{t-1}, and b_t. The latter property suggests that all warehouse firms hold the same amount of inventory at $p_t = P_L$. At $p_t = P_H$, equation (38) suggests that i_H is increasing in $r(I_t, P_H, b)$ and decreasing in $r(I_{t-1}, P_H, b_t)$. Furthermore, equation (38) leads to $\mu_t - C'(i_H) = -\beta \left\{(1 - \overline{q}) r(I_t, P_H, b) + \overline{q} P_L\right\} + r(I_{t-1}, P_H, b_t)$, which suggest that $\mu_t = 0$ and $i_H > 0$ if $\beta \left\{(1 - \overline{q}) r(I_t, P_H, b) + \overline{q} P_L\right\} > r(I_{t-1}, P_H, b_t)$. This means that, if the price to be sold to the household is expected to increase much, warehouse firms hold some amount of inventory at $p_t = P_H$. Otherwise, warehouse firms do not hold inventory, that is, $i_H = 0$ at $p_t = P_H$. The inventory i_H is decreasing in $r(I_{t-1}, P_H, b_t)$. If $I_{t-1} = 0$, then $r(I_{t-1}, P_H, b_t) = P_H$. The household purchases directly from producers.

Equations (37) and (38) suggest that, if $i_t > 0$, expected $r(I_t, P_H, b)$ should increase from $r(I_{t-1}, P_H, b)$, because otherwise the right-hand side of the equations is negative, while the left-hand sige is positive. Since I_t decreases over time, this suggests that $r(I_t, P_H, b)$ is decreasing in I_t.

Finally, $r(I_{t-1}, P_H, b_t)$ is increasing in b_t because household demand is increasing in b_t according to equation (15). At $p_t = P_L$, $r(I_{t-1}, P_L, b_t)$ is independent of b_t unless a big negative shock of b_t induces fire sales so that $r(I_{t-1}, P_L, b_t)$ is increasing in b_t when $b_t \ll 0$.

Equation (35) suggests that $\psi_t > 0$, that is, $x_t = 0$, when $r_t < p_t$. In other words, warehouse firms do not purhcase goods ($x_t = 0$) if $p_t = P_H$. ■

Proof of Lemma 3

Proof. Consider a sales period from $t+1$ to $t+T$ ($T \geq 2$). For the former statement, we examine the case of $p_t = p_{t+T+1} = P_H$. Then, from the previous lemma, warehouse firms do not purchase, i.e., $x_t = x_{t+T+1} = 0$. Because the inventory at P_L, i_L, is independent of i_{t-1} and I_{t-1} and inventory at P_H, i_H, is decreasing over time, we should see $I_{t-1} \leq I_{t+T}$. Thus, $r_H(I_{t-1}) \geq r_H(I_{t+T})$. Because households do not purchase goods from producers when $r_t < P_H$, the quantity purchased by households is greater than or equal to that just after sales.

For the latter statement, consider a sales period from $t+1$ to $t+T$ ($T \geq 3$). It is obvious to see $I_t \leq I_{t+T-1}$. Because i_L is independent of i_{t-1} and I_{t-1},the quantity purchased by warehouse firms at the first day of sales is greater than or equal to that at the final day of sales. I_{t-1},The quantity purchased by households is the same between the first and last days of sales, because household face the price of P_L. ■

Proof of Lemma 4

Proof. We showed that the inventory i_H at $p_t = P_H$ is decreasing over time in the proof of the previous lemma. Thus, in aggregate I_t is decreasing over time after sales end, untile I_t reaches zero. Since $r_t = r(I_{t-1}, p_t, b_t)$ is non-increasing in I_{t-1}, r_t is non-decreasing over time. Because of equation (15), c_t is decreasing with r_t, and thus non-increasing over time. ■

Proof of Lemma 5

Proof. From equation (25), the COLI in each period is written as

$$\mathcal{C}(r_t) = C(r_{t+T+1+T_H}) = \left[\left(b^k\right)^{\sigma} (P_H)^{1-\sigma} U + A\right]^{1/(1-\sigma)}$$

where A is some constant. Thus, it is immediate to see that $\pi^{COLI} = \log\left\{\Pi_{i=1}^{T+1+T_H}(\mathcal{C}(r_{t+i})/\mathcal{C}(r_{t+i-1}))\right\}$ equals zero.

The total quantity puchased by warehouse firms and households at t equals $X_t = \int_0^{N_t} x_t dj + \int_0^{M_t} z_t dj$. Regarding this, we showed in Lemma 3 that $X_t \geq X_{t+T+1}$ and $X_{t+1} \geq X_{t+T}$. Further, if $m \geq 1$, the inventory at the end of sales is not zero but positive, so $X_t > X_{t+T+1}$ and $X_{t+1} > X_{t+T}$. Then, using $\sum_{k' \in K_0 \cap K_t} p_t^{k'} x_t^{k'} = 1$,we obtain

the purchase-weighted Törnqvist index as

$$\begin{aligned}\pi^T = \sum_{i=1}^{T+1+T_H} \pi_{t+i}^T &= \frac{p_t X_t + p_{t+1} X_{t+1}}{2} \log\left(\frac{p_{t+1}}{p_t}\right) + \frac{p_{t+T} X_{t+T} + p_{t+T+1} X_{t+T+1}}{2} \log\left(\frac{p_{t+T+1}}{p_{t+T}}\right) \\ &= \frac{P_H X_t + P_L X_{t+1}}{2} \log\left(\frac{P_L}{P_H}\right) + \frac{P_L X_{t+T} + P_H X_{t+T+1}}{2} \log\left(\frac{P_H}{P_L}\right) \\ &= \frac{P_L \left(X_{t+T} - X_{t+1}\right) + P_H \left(X_{t+T+1} - X_t\right)}{2} \log\left(\frac{P_H}{P_L}\right) \\ &< 0.\end{aligned}$$

Similarly, we have

$$\begin{aligned}\pi^L = \sum_{i=1}^{T+1+T_H} \pi_{t+i}^L &= p_t X_t \log\left(\frac{p_{t+1}}{p_t}\right) + p_{t+T} X_{t+T} \log\left(\frac{p_{t+T+1}}{p_{t+T}}\right) \\ &= (P_L X_{t+T} - P_H X_t) \log\left(\frac{P_H}{P_L}\right) \\ &\propto (P_L P_L^{-\sigma} - P_H P_H^{-\sigma}) \log\left(\frac{P_H}{P_L}\right) \\ &< (>) 0 \quad \text{if } \sigma > (<) 1,\end{aligned}$$

where $A \propto B$ means that A is positively proportional to B. Note that demand is proportional to $p_t^{-\sigma}$ at t and $t+T$ because there is no demand for further stockpiling. Also, we have

$$\begin{aligned}\pi^P = \sum_{i=1}^{T+1+T_H} \pi_{t+i}^P &= p_{t+1} X_{t+1} \log\left(\frac{p_{t+1}}{p_t}\right) + p_{t+T+1} X_{t+T+1} \log\left(\frac{p_{t+T+1}}{p_{t+T}}\right) \\ &= (P_H X_{t+T+1} - P_L X_{t+1}) \log\left(\frac{P_H}{P_L}\right) \\ &< \frac{P_L \left(X_{t+T} - X_{t+1}\right) + P_H \left(X_{t+T+1} - X_t\right)}{2} \log\left(\frac{P_H}{P_L}\right) \\ &= \pi^T \quad \text{if } \sigma > 1,\end{aligned}$$

because $P_H \left(X_{t+T+1} + X_t\right) < 2P_H X_t < 2P_L X_{t+T} < P_L \left(X_{t+T} + X_{t+1}\right)$.

If $m \geq 1$, the effective price satisfies $P_L < r_{t+T+1} < P_H$ just after sales and $r_{t+T+j+1} \geq r_{t+T+j}$ for $j = 1, 2, \cdots, T_H$. Also, we have $r_t = P_H$ (i.e. just before sales) and $r_{t+j} = P_L$ for $j = 1, \cdots, T$ (i.e. during sales).

We obtain the consumption-weighted Törnqvist index as

$$\begin{aligned}\pi^{T*} = \sum_{i=1}^{T+1+T_H} \pi_{t+i}^{T*} &= \frac{r_t c_t + r_{t+1} c_{t+1}}{2} \log\left(\frac{r_{t+1}}{r_t}\right) + \sum_{i=T+1}^{T+1+T_H} \frac{r_{t+i-1} c_{t+i-1} + r_{t+i} c_{t+i}}{2} \log\left(\frac{r_{t+i}}{r_{t+i-1}}\right) \\ &= -\frac{P_H c_t + P_L c_{t+1}}{2} \log\left(\frac{P_H}{P_L}\right) + \sum_{i=T+1}^{T+1+T_H} \frac{r_{t+i-1} c_{t+i-1} + r_{t+i} c_{t+i}}{2} \log\left(\frac{r_{t+i}}{r_{t+i-1}}\right).\end{aligned}$$

The effective price drops from P_H to P_L in one period and increases to P_H taking periods

longer than one. If it takes two periods, π^{T*} is

$$\begin{aligned}
\pi^{T*} = \sum_{i=1}^{T+1+T_H} \pi_{t+i}^{T*} &= -\frac{P_H c_t + P_L c_{t+1}}{2}\log\left(\frac{P_H}{P_L}\right) + \sum_{i=T+1}^{T+2} \frac{r_{t+i-1}c_{t+i-1} + r_{t+i}c_{t+i}}{2}\log\left(\frac{r_{t+i}}{r_{t+i-1}}\right) \\
&\propto -\frac{P_H P_H^{-\sigma} + P_L P_L^{-\sigma}}{2}\log\left(\frac{P_H}{P_L}\right) + \frac{P_L P_L^{-\sigma} + r_H r_H^{-\sigma}}{2}\log\left(\frac{r_H}{P_L}\right) + \frac{r_H r_H^{-\sigma} + P_H P_H^{-\sigma}}{2}\log\left(\frac{P_H}{r_H}\right) \\
&= -\frac{P_H P_H^{-\sigma} + P_L P_L^{-\sigma}}{2}\log\left(\frac{P_H}{P_L}\right) + \frac{P_L P_L^{-\sigma} + r_H r_H^{-\sigma}}{2}\log\left(\frac{P_H}{P_L}\right) \\
&\quad -\frac{P_L P_L^{-\sigma} + r_H r_H^{-\sigma}}{2}\log\left(\frac{P_H}{r_H}\right) + \frac{r_H r_H^{-\sigma} + P_H P_H^{-\sigma}}{2}\log\left(\frac{P_H}{r_H}\right) \\
&= \frac{r_H^{1-\sigma} - P_H^{1-\sigma}}{2}\log\left(\frac{P_H}{P_L}\right) - \frac{P_L^{1-\sigma} - P_H^{1-\sigma}}{2}\log\left(\frac{P_H}{r_H}\right) \\
&\equiv F(r_H).
\end{aligned}$$

This equals zero when $r_H = P_H$ or $r_H = P_L$.

The function $F(r_H)$ suggests that $2r_H^{\sigma}F'(r_H) = (r_H/P_L)^{\sigma-1} - (r_H/P_H)^{\sigma-1} - (\sigma - 1)\log(P_H/P_L)$. Suppose $\sigma > 1$. This is increasing in r_H, negative when $r_H = P_L$, and positive when $r_H = P_H$. This means that there exits a certain $r_H^* \subset (P_L, P_H)$ that makes $F'(r_H^*)$ equal to zero. Furthermore, $F'(r_H) < 0$ when $P_L \leq r_H < r_H^*$ and $F'(r_H) > 0$ when $r_H^* \leq r_H < P_H$. Thus, $F(r_H)$ is negative when $P_L < r_H < P_H$. Although what we showed here is limited to the case in which the effective price decreases and increases in one period and two periods, respectively, this logic should also hold in the case in which the effective price decreases and increases in one period and periods longer than two, respectively. For example, suppose an increase taking three periods such that $r_{t+T} = P_L < r_{t+T+1} < r_{t+T+2} < r_{t+T+3} = P_H$. If we hypothetically divide a decrease in the effective price from t to $t+1$ to P_H to r_{t+T+2} and r_{t+T+2} to P_L,we should have $\pi^{T*} = 0$ for the price change between r_{t+T+2} and P_H, but $\pi^{T*} < 0$ for the price change between P_L and r_{t+T+2} because the price increase takes two periods. Thus, in sum, we should have $\pi^{T*} < 0$.

On the other hand, if $\sigma < 1$, $F(r_H)$ is positive when $P_L < r_H < P_H$,which leads to $\pi^{T*} > 0$.

We compare π^{T*} and π^{T} in the special case of $m=1$:

$$\begin{aligned}
\pi^{T*}-\pi^{T} \propto & \frac{r_H^{1-\sigma}-P_H^{1-\sigma}}{2}\log\left(\frac{P_H}{P_L}\right)-\frac{P_L^{1-\sigma}-P_H^{1-\sigma}}{2}\log\left(\frac{P_H}{r_H}\right) \\
& -\frac{P_L\left(X_{t+T}-X_{t+1}\right)+P_H\left(X_{t+T+1}-X_t\right)}{2}\log\left(\frac{P_H}{P_L}\right) \\
= & \frac{r_H^{1-\sigma}-P_H^{1-\sigma}}{2}\log\left(\frac{P_H}{P_L}\right)-\frac{P_L^{1-\sigma}-P_H^{1-\sigma}}{2}\log\left(\frac{P_H}{r_H}\right) \\
& -\frac{P_L I_L+P_H\left(0-P_H^{-\sigma}\right)}{2}\log\left(\frac{P_H}{P_L}\right) \\
= & \frac{r_H^{1-\sigma}+P_L I_L}{2}\log\left(\frac{P_H}{P_L}\right)-\frac{P_L^{1-\sigma}-P_H^{1-\sigma}}{2}\log\left(\frac{P_H}{r_H}\right) \\
\equiv & G(r_H).
\end{aligned}$$

Suppose $\sigma>1$. As we showed in the above proof, $G'(r_H)$ takes the smallest value at $r_H=r_H^*$, where r_H^* satisfies $\left(r_H^*/P_L\right)^{\sigma-1}-\left(r_H^*/P_H\right)^{\sigma-1}-(\sigma-1)\log\left(P_H/P_L\right)$. Substituting r_H^* into the above equation, we can find that $G(r_H^*)$ is positive. Thus, $G(r_H)$ is positive. ■

Proof of Lemma 6

Proof. Using Shephard's Lemma and equation (27), we have

$$\begin{aligned}
c_t^i &= U_t\partial\mathcal{C}(r_t)/\partial r_t^i \\
&= U_t\frac{1}{2(1-\sigma)}\left[\sum_{i\in K}\sum_{k\in K}\alpha^{ik}\left(r_t^i\right)^{(1-\sigma)}\left(r_t^k\right)^{(1-\sigma)}\right]^{1/\{2(1-\sigma)\}-1} 2(1-\sigma)\left(r_t^i\right)^{-\sigma}\sum_{k\in K}\alpha^{ik}\left(r_t^k\right)^{1-\sigma} \\
&= \frac{U_t\mathcal{C}(r_t)\left(r_t^i\right)^{-\sigma}\sum_{k\in K}\alpha^{ik}\left(r_t^k\right)^{1-\sigma}}{\sum_{i\in K}\sum_{k\in K}\alpha^{ik}\left(r_t^i\right)^{(1-\sigma)}\left(r_t^k\right)^{(1-\sigma)}},
\end{aligned}$$

which leads to

$$\frac{r_t^i c_t^i}{U_t\mathcal{C}(r_t)}=\frac{\left(r_t^i\right)^{1-\sigma}\sum_{k\in K}\alpha^{ik}\left(r_t^k\right)^{1-\sigma}}{\sum_{i\in K}\sum_{k\in K}\alpha^{ik}\left(r_t^i\right)^{(1-\sigma)}\left(r_t^k\right)^{(1-\sigma)}}. \tag{39}$$

Noting

$$\begin{aligned}
s_t^i &\equiv \frac{r_t^i c_t^i}{\sum_{k\in K} r_t^k c_t^k} \\
&= \frac{r_t^i c_t^i}{U_t\mathcal{C}(r_t)} \\
&= \frac{\left(r_t^i\right)^{1-\sigma}\sum_{k\in K}\alpha^{ik}\left(r_t^k\right)^{1-\sigma}}{\sum_{i\in K}\sum_{k\in K}\alpha^{ik}\left(r_t^i\right)^{(1-\sigma)}\left(r_t^k\right)^{(1-\sigma)}},
\end{aligned} \tag{40}$$

we have

$$
\begin{aligned}
P_r(r_0, r_1, c_0, c_1) &= \frac{\left\{\sum_{i\in K} s_0^i \left(\frac{r_1^i}{r_0^i}\right)^{(1-\sigma)}\right\}^{1/\{2(1-\sigma)\}}}{\left\{\sum_{k\in K} s_1^i \left(\frac{r_0^i}{r_1^i}\right)^{(1-\sigma)}\right\}^{1/\{2(1-\sigma)\}}} \\
&= \frac{\left\{\sum_{i\in K} \frac{\left(r_0^i\right)^{1-\sigma}\sum_{k\in K}\alpha^{ik}\left(r_0^k\right)^{1-\sigma}}{\sum_{i\in K}\sum_{k\in K}\alpha^{ik}\left(r_0^i\right)^{(1-\sigma)}\left(r_0^k\right)^{(1-\sigma)}} \left(\frac{r_1^i}{r_0^i}\right)^{(1-\sigma)}\right\}^{1/\{2(1-\sigma)\}}}{\left\{\sum_{i\in K} \frac{\left(r_1^i\right)^{1-\sigma}\sum_{k\in K}\alpha^{ik}\left(r_1^k\right)^{1-\sigma}}{\sum_{i\in K}\sum_{k\in K}\alpha^{ik}\left(r_1^i\right)^{(1-\sigma)}\left(r_1^k\right)^{(1-\sigma)}} \left(\frac{r_0^i}{r_1^i}\right)^{(1-\sigma)}\right\}^{1/\{2(1-\sigma)\}}} \\
&= \frac{\left\{\frac{\sum_{i\in K}\sum_{k\in K}\alpha^{ik}\left(r_1^i\right)^{(1-\sigma)}\left(r_0^k\right)^{(1-\sigma)}}{\sum_{i\in K}\sum_{k\in K}\alpha^{ik}\left(r_0^i\right)^{(1-\sigma)}\left(r_0^k\right)^{(1-\sigma)}}\right\}^{1/\{2(1-\sigma)\}}}{\left\{\frac{\sum_{i\in K}\sum_{k\in K}\alpha^{ik}\left(r_0^i\right)^{(1-\sigma)}\left(r_1^k\right)^{(1-\sigma)}}{\sum_{i\in K}\sum_{k\in K}\alpha^{ik}\left(r_1^i\right)^{(1-\sigma)}\left(r_1^k\right)^{(1-\sigma)}}\right\}^{1/\{2(1-\sigma)\}}} \\
&= \left\{\frac{\sum_{i\in K}\sum_{k\in K}\alpha^{ik}\left(r_1^i\right)^{(1-\sigma)}\left(r_1^k\right)^{(1-\sigma)}}{\sum_{i\in K}\sum_{k\in K}\alpha^{ik}\left(r_0^i\right)^{(1-\sigma)}\left(r_0^k\right)^{(1-\sigma)}}\right\}^{1/\{2(1-\sigma)\}} \\
&= \frac{\mathcal{C}(r_1)}{\mathcal{C}(r_0)}.
\end{aligned}
$$

■

Equilibrium When Inventory is Cleared in Just One Period After Sales

The warehouse firm's value is written as

$$V(i_{t-1}, p_t; I_{t-1}) = \max_{x_t, i_t}\left(r_t(i_{t-1} - i_t + x_t) - p_t x_t - C(i_t)\right) + \beta \mathrm{E}_t\left[V(i_t, p_{t+1}; I_t)\right].$$

The free entry condition leads to

$$V(i_{t-1} = 0, p_t = P_L; I_{t-1}) = 0, \tag{41}$$

while $V(i_{t-1} = 0, p_t = P_H; I_{t-1}) < 0$. This condition determines r_H as a function of $r_H(I_{t-1})$.

If warehouse firms do not hold much inventory, they sell off all their inventory immediately when p_t turns from P_L and P_H. That is, i_H becomes zero in the next period. The firm value can be divided into four cases:

$$
\begin{aligned}
V(i_L, P_L; I_L) &= -C(i_L) + \beta\left\{(1-\underline{q})V(i_L, P_H; I_L) + \underline{q}V(i_L, P_L; I_L)\right\}, \\
V(i_L, P_H; I_L) &= r_H(I_L)i_L - C(0) + \beta\left\{(1-\overline{q})V(0, P_H; 0) + \overline{q}V(0, P_L; 0)\right\}, \\
V(0, P_L; 0) &= -P_L i_L - C(i_L) + \beta\left\{(1-\underline{q})V(i_L, P_H; I_L) + \underline{q}V(i_L, P_L; I_L)\right\} = 0, \\
V(0, P_H; 0) &= -C(0) + \beta\left\{(1-\overline{q})V(0, P_H; 0) + \overline{q}V(0, P_L; 0)\right\}.
\end{aligned}
$$

These equations are summarized as

$$C(i_L) + (1 - \beta\underline{q})P_L i_L + \frac{\beta(1-\underline{q})}{1-\beta(1-\overline{q})}C(0) = \beta(1-\underline{q})r_H(I_L)i_L.$$

Because equation (37) is rewritten as

$$C'(i_L) + (1 - \beta\underline{q})P_L = \beta(1-\underline{q})r_H(I_L), \tag{42}$$

we have

$$C'(i_L)i_L - C(i_L) = \frac{\beta(1-\underline{q})}{1-\beta(1-\overline{q})}C(0). \tag{43}$$

Equations (42) and (43) enable us to solve for i_L and $r_H(I_L)$.

Finally, it is important to check that the above solution satisfies the goods market clearing condition. When $p_t = P_L$, incumbent firms supply goods with no additional cost until the goods market is cleared, because output price r_t is equal to input price P_L. In the case of $p_t = p_{t-1} = P_H$, we have $r_t = P_H$. The household purchases storable goods k directly from a goods k producer until the goods market is cleared. In the case of $p_t = P_H$ and $p_{t-1} = P_L$, $r_t = r_H$ and the quantity of goods supply is predetermined because the inventory is perdetermined and there is no additional purchase $x_t = 0$. Thus, if demand is greater (smaller) than supply in the goods market, r_H increases (decreases) until the market is cleared, which makes the firm value deviates from its expected value. In other words, the expected price of r_{t+1} conditional on $p_t = P_H$ and $p_{t-1} = P_L$ equals r_H, but the realized price may differ from the expected value. Also, it is worth pointing out that r_H is independent of factors at the household side, that is, b_t^k. They influence demand for storable goods k, but do not influence the price r_H because firms freely enter the market and change aggregate goods supply.

C Detailed Explanations about Our Approach

Proof of m_{cont} Suppose continuous time. If r_t is linearly increasing with t, we can write $r_H(I_t)$ as $r_H(x) = \left\{\frac{x}{m_{cont}}\frac{P_H - P_L}{P_L} + 1\right\}P_L$, where x represents the time elasped after sales, because $r_H(0) = P_L$ and $r_H(m_{cont}) = P_H$. Corresponding consumption equals $(r_H/P_L)^{-\sigma} c_L^*$. At $m = m_{cont}$, all inventory is used, which is given by

$$I_L = \int_0^m \left\{\frac{x}{m_{cont}}\frac{P_H - P_L}{P_L} + 1\right\}^{-\sigma} c_L^* dx$$
$$\therefore m_{cont} = \frac{P_H - P_L}{P_L}\frac{\sigma - 1}{1 - (P_H/P_L)^{-\sigma+1}}\frac{I_L}{c_L^*}.$$

Condition of Our Approach In equation (28), $r_H(I_t) - r_H(I_{t-1})$ is positive constant if

$$C'(i_H; I_{t-1}) + \{1 - \beta(1 - \overline{q})\} r_H(I_{t-1})$$

is constant and larger than $\beta \overline{q} P_L$.

In particular, the value just after sales should equal that when inventory disappears:

$$C'(\kappa I_L) + \{1 - \beta(1 - \overline{q})\} P_L = C'(0) + \{1 - \beta(1 - \overline{q})\} P_H > \beta \overline{q} P_L,$$

where $i_H = \kappa I_t$ in equilibrium with a positive coefficient κ. Thus, we have

$$C'(\kappa I_L) - C'(0) = \{1 - \beta(1 - \overline{q})\} (P_H - P_L) > 0 \tag{44}$$

and

$$C'(0) > \beta \overline{q} P_L - \{1 - \beta(1 - \overline{q})\} P_H. \tag{45}$$

This suggests that C' is increasing (i.e., $C'' > 0$) and $C'(0)$ is sufficiently large.

More generally, the condition is given as follows. Note that, if $r_H(I_t) - r_H(I_{t-1})$ is positive constant, then the effective price and inventory at x ($0 \leq x \leq m_{cont}$) equal

$$\begin{aligned} r_H(x) &= \left\{ \frac{x}{m_{cont}} \frac{P_H - P_L}{P_L} + 1 \right\} P_L, \\ I(x) &= I_L - \int_0^x \left\{ \frac{x'}{m_{cont}} \frac{P_H - P_L}{P_L} + 1 \right\}^{-\sigma} c_L^* dx' \\ &= \left\{ 1 - \frac{1 - \left(\frac{x}{m_{cont}} \frac{P_H - P_L}{P_L} + 1 \right)^{-\sigma+1}}{1 - (P_H/P_L)^{-\sigma+1}} \right\} I_L, \end{aligned}$$

respectively. Thus, inserting $i_H = \kappa I_t$, we can rewrite the above equation as

$$C' \left(\left\{ 1 - \frac{1 - \left(\frac{x}{m_{cont}} \frac{P_H - P_L}{P_L} + 1 \right)^{-\sigma+1}}{1 - (P_H/P_L)^{-\sigma+1}} \right\} \kappa I_L \right) + \{1 - \beta(1 - \overline{q})\} \left\{ \frac{x}{m_{cont}} \frac{P_H - P_L}{P_L} + 1 \right\} P_L, \tag{46}$$

which needs to be independent of x and and larger than $\beta \overline{q} P_L$. Defining

$$x' \equiv \left\{ 1 - \frac{1 - \left(\frac{x}{m_{cont}} \frac{P_H - P_L}{P_L} + 1 \right)^{-\sigma+1}}{1 - (P_H/P_L)^{-\sigma+1}} \right\} \kappa I_L$$

for $0 \leq x' \leq \kappa I_L$, we have

$$C'(x') = C - \{1 - \beta(1 - \overline{q})\} \left\{ 1 - \left(1 - \frac{x'}{\kappa I_L} \right) \left(1 - \left(\frac{P_H}{P_L} \right)^{-\sigma+1} \right) \right\}^{\frac{1}{-\sigma+1}} P_L, \tag{47}$$

where $C > \beta \overline{q} P_L$. It can be easily shown that $C'''(x') > 0$.

Details of Our Approach when Applying to the POS Data To infer $r_H(I_{t-1})$ and m, we need to have σ, c_L^*, and I_L from the POS data. More precisely, we calculate $r_H(I_{t-1})$, m, X_L^*, and I_L for each sales event of product k and shop s, while we calculate σ for each 3-digit product category.

We take the following three steps. First, for product k and shop s, we identify all sales events using the V-shape filter. Define the first and last days of sales by $t+1$ and $t+T$, respectively. We record the price and the quantity purchased just before sales (i.e., at t) as r_H and c_H, respectively. Note that, according to the model, c_H should equal to consumption, unless the effect of stockpiling remains and c_H is zero. If so, we use the previous values of r_H and c_H.

Second, if sales last more than one day ($T > 1$), we record the average price and the quantity purchased in the first half spell of sales (i.e. from $t+1$ to $t+\lfloor T/2 \rfloor - 1$) as P_L^1 and X_L^1, while the average price and the quantity purchased in the second half spell of sales (i.e. from $t+\lfloor T/2 \rfloor$ to $t+T$) are defined as P_L^2 and X_L^2, respectively. If $X_L^1 \geq X_L^2$, we set $r_L = P_L^2$ and $c_L = X_L^2$. Otherwise, we set $r_L = P_L^1$ and $c_L = X_L^1$. According to the model, we should observe $X_L^1 \geq X_L^2$ as Lemma 3 showed. In such a case, there is no need for additional stockpiling in the second half spell of sales, so X_L^2 should equal consumption, which we denote by c_L. However, the POS data often show the opposite: $X_L^1 < X_L^2$, which is justifiable if the duration of sales is known *ex ante*. In such a case, there is no need for stockpiling in the first half spell of sales, so X_L^1 should equal c_L.

In the third step, we calculate the elasticity of substitution σ. The variables we obtained in Steps 1 and 2, r_H, r_L, c_H and c_L, correspond to the effective prices and cosumption. Thus, equation (15) should hold among these variables. Furthermore, the log ratio of quantity consumed at sales to that at regular divided by the log ratio of price at sales to that at regular, $\Gamma \equiv -\log(c_L/c_H)/\log(r_L/r_H)$, should equal σ on average. We calculate the unweighted average of Γ across sales events, products and shops for each 3-digit product category, which we define as σ.

Fourth, we calculate consumption, the effective price, the degree of stockpiling, and inventory for each sales event of product k and shop s. If sales last more than one day ($T > 1$), we set $c_L^* = (r_L/r_H)^{-\sigma} c_H$. If sales are one day long, we set $r_L = p_t$ and $c_L^* = \min\left[0.01, (r_L/r_H)^{-\sigma} c_H\right]$. Inventory at the end of sales, I_L, equals the cumulated quantity of purchase during sales minus the cumulated quantity of consumption during sales, that is, $\max\left[0, \Sigma_{j=1}^{T} X_{t+j} - T c_L^*\right]$, where we replace it with zero if calculated inventory is negative. Once we obtain c_L^* and I_L, we can calculate m. Note that, since the unit of

time is discrete (a day) in our analysis, we search for the maximum natural number m starting from the above continuous-time version of m_{cont} so that the inventory at $t+T+m$ is positive, while that at $t+T+m+1$ is negative. Once we obtain m, we can calculate the path of effective prices $r_s = r_H(I_{s-1})$ and consumption c_s as $r_L+((s-t-T)/m)(r_H-r_L)$ and $(r_s/r_L)^{-\sigma} c_L^*$, respectively, for $t+T+1 \leq s \leq t+T+m$.[16]

[16]There is a possibility that the next sales begin at $t^{'}$ before $t+T+m$. In such a case, we use r_s and c_s until $t^{'}-1$ and recalculate them for $s \geq t^{'}$ assuming that new sales start with no inventory ($I_{t^{'}-1} = 0$).

Table 1: Backgrounds of the Chain Drift

		$t=1$	$t=2$	$t=3$
Product A	Price	p_A	$(1-r)p_A$	p_A
	Share	W_1	W_2	W_3
Product B	Price	p_B	p_B	p_B
	Share	$1-W_1$	$1-W_2$	$1-W_3$

Note: $0 < r < 1$. If the elasticity of substitutions is greater than one, it is natural to observe $W_3 < W_1 < W_2$.

Table 2: State-dependent Consumption

Dependent variable	1 if used	No. of uses
Inventory	0.0006895***	0.00410***
	(0.0000171)	(0.000901)
Observations	90,545,020	90,545,020
No. of HH	3,602	3,602
R^2	0.12	0.12
Fixed effects	HH/category	HH/category

Note: Figures in parentheses represent robust standard errors. ***, **, and * represent significance at the 1%, 5%, and 10% levels, respectively.

Table 3: Simulation of the Chain Drift

	COLI	Quadratic mean of r	Tornqvist (C)	Tornqvist (Q)	Laspeyres (Q)	Paasche (Q)
Benchmark	1.000	1.000	0.991	0.163	2.71e+01	1.01e-03
	(4.60e-03)	(4.60e-03)	(4.58e-03)	(1.09e-02)	(3.80e+00)	(2.10e-04)
Low m	1.000	1.000	1.000	0.448	2.73e+00	7.36e-02
$(m = 1)$	(3.88e-03)	(3.88e-03)	(3.91e-03)	(1.15e-02)	(1.02e-01)	(5.82e-03)
High m	0.999	0.999	0.992	0.054	3.42e+02	9.13e-06
$(m = 10)$	(5.08e-03)	(5.08e-03)	(5.06e-03)	(6.20e-03)	(8.82e+01)	(3.15e-06)
Low σ	1.000	1.000	0.999	0.177	1.71e+01	1.87e-03
$(\sigma = 2)$	(4.29e-03)	(4.29e-03)	(4.29e-03)	(1.14e-02)	(2.16e+00)	(3.63e-04)
Low P_L/P_H	0.999	0.999	0.808	0.005	2.50e+05	1.44e-10
$(P_L/P_H = 0.75)$	(1.48e-02)	(1.48e-02)	(1.33e-02)	(1.01e-03)	(1.25e+05)	(9.40e-11)
High $\overline{q}$	1.001	1.001	0.980	0.005	1.87e+05	1.46e-10
$(\overline{q} = 1/7)$	(6.06e-03)	(6.06e-03)	(5.94e-03)	(3.50e-04)	(4.24e+04)	(2.59e-11)
High $\underline{q}$	1.000	1.000	0.991	0.163	2.71e+01	1.01e-03
$(\underline{q} = 0.5)$	(4.60e-03)	(4.60e-03)	(4.58e-03)	(1.09e-02)	(3.80e+00)	(2.10e-04)

Note: Figures indicate the mean of price levels after 365 days, where initial price levels are set at one. Figures in parentheses indicate the standard deviation.

Table 4: Regression of the Chain Drift

	Coef	SE
log(m)	-0.0013***	(0.0005)
$\overline{q}$	-0.0114***	(0.0043)
$\underline{q}$	0.0015	(0.0013)
$\log(P_L/P_H)$	0.007***	(0.0022)
constant	-0.0003	(0.0005)
Adjusted R^2	0.348	
Observations	145	

Note: The dependent variable is the daily average of the inflation rate from January 1989 to December 2011 based on the purchase-weighted Törnqvist index for each 3-digit product category. ***, **, and * represent significance at the 1%, 5%, and 10% levels, respectively.

Table 5: Top XX Categories for the Degree of Stockpiling, Inventory Duration, and the Quantiry of Purchase

Ranking of m and inventory

Note: The degree of stockpiling m is inferred using the POS data. The inventory duration and the quantiry of purchase are taken from the Shokumap data.

Table 6: Regression of Stockpiling

$\log(m)$	(1)	(2)
$\log(1-\overline{q})$	1.7463***	-0.901
	(0.542)	(0.741)
$\log(1-\underline{q})$	-0.6306***	-0.8058***
	(0.041)	(0.038)
$\log(1-P_L/P_H)$	0.3897***	0.3349***
	(0.047)	(0.045)
Fixed effects	category	category/month
Observations	40,296	40,296
R^2	0.588	0.719
# of categories	146	146

Note: ***, **, and * represent significance at the 1%, 5%, and 10% levels, respectively.

Table 7: Business Cycle Dependence of Stockpiling

$\log(m)$	(1)	(2)	(3)	(4)	(5)
Unemployment rate	-0.1379***		-0.0171***	-0.0252***	-0.0382***
	(0.009)		(0.005)	(0.004)	(0.004)
log(hours worked)		4.5572***	4.1587***	2.1906***	2.8795***
		(0.281)	(0.230)	(0.155)	(0.134)
$\log(1-\overline{q})$				-3.0092***	-1.3061*
				(0.547)	(0.675)
$\log(1-\underline{q})$				-0.6092***	-0.6766***
				(0.036)	(0.037)
$\log(1-P_L/P_H)$				0.3297***	0.3515***
				(0.043)	(0.041)
Fixed effects	category	category	category	category	category
time trend					yes
Observations	40,296	40,296	40,296	40,296	40,296
R^2	0.258	0.351	0.353	0.650	0.657
# of categories	146	146	146	146	146

Note: ***, **, and * represent significance at the 1%, 5%, and 10% levels, respectively.

Figure 1: Chain Drift

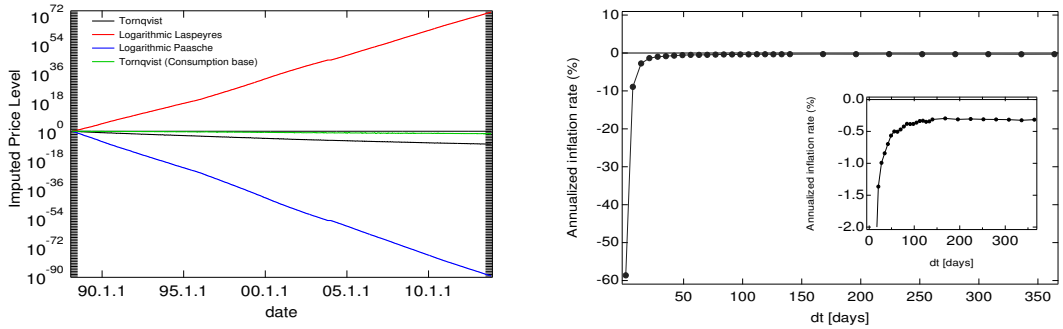


Figure 2: Consumption Pattern of Salt

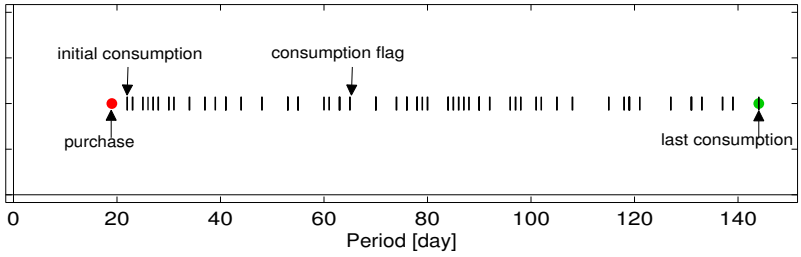


Figure 3: Purchase and Consumption Patterns

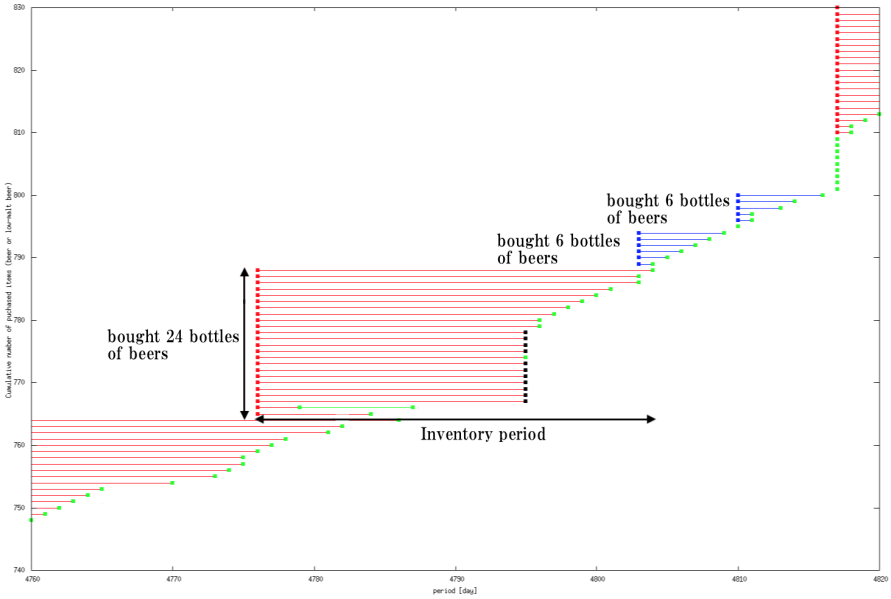
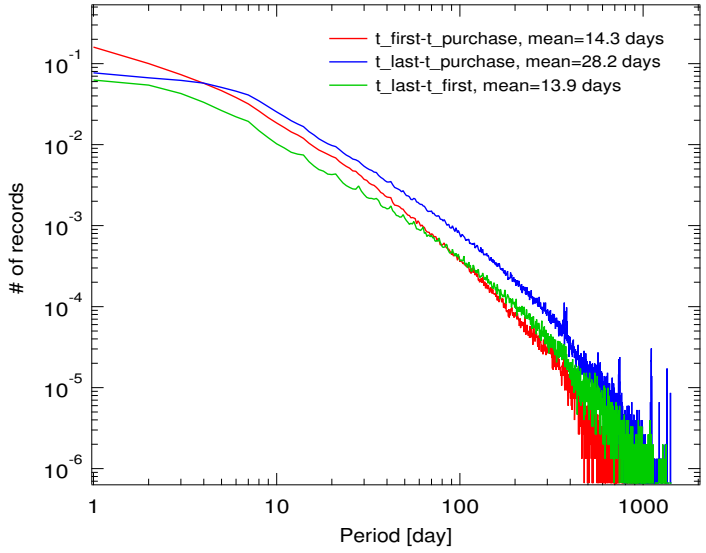


Figure 4: Density of Consumption Duration



Note: This shows the density of consumption durations in the Shoku-map data. We count the durations among the date of puchase (t_p), the date of the first consumption (t_f), and the date of the last consumption (t_l), for each product i and household j.

Figure 5: Changes in the Price Index ($dt = 365$)

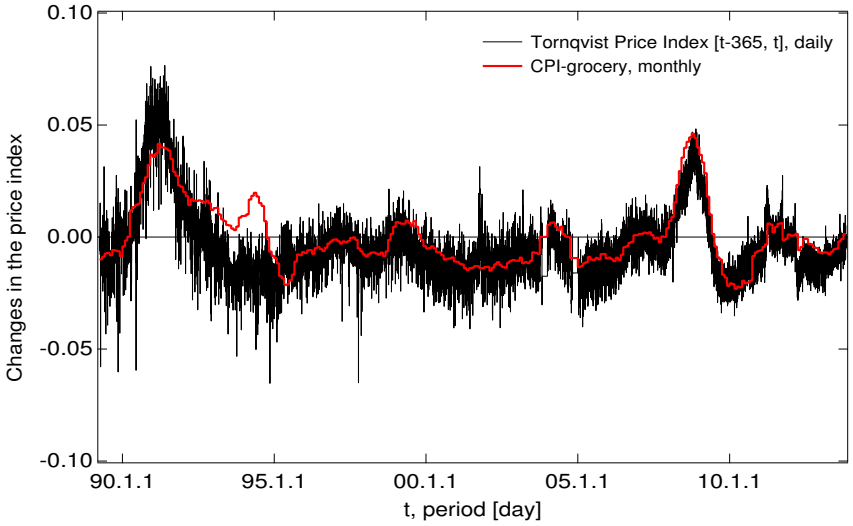
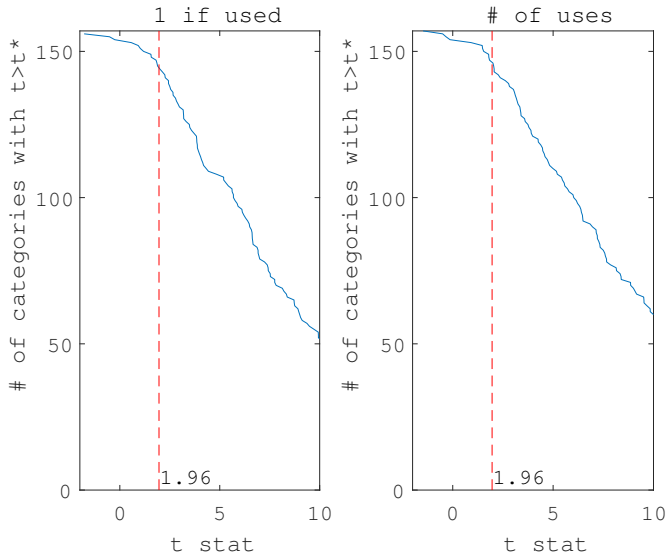
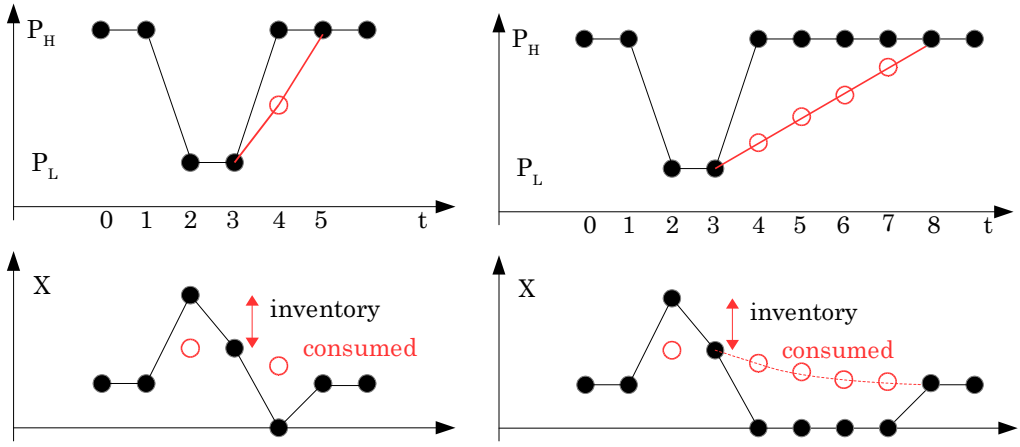


Figure 6: Distribution for the Estimate of the Inventory Elasticity of Consumption



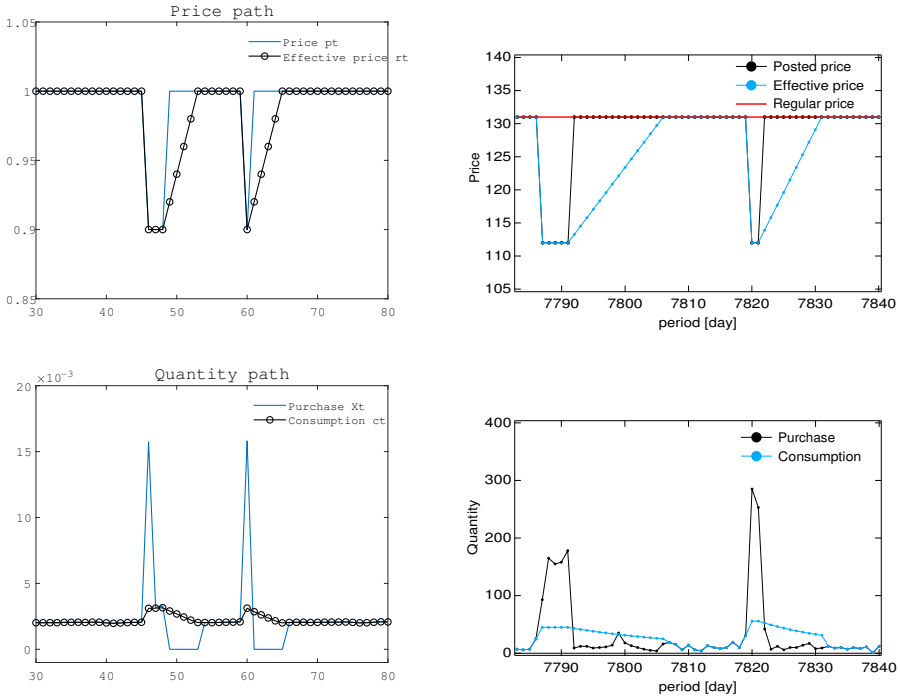
Note: We regress consumption for each 3-digit category level and obtain the t statistics for the coefficient on inventory. The vertical axis indicates the number of categories whose t statistics for the inventory elasticity of consumption is larger than t^*, while the horizontal axis represents t^*. The dependent variable is the consumption that takes one if used for the left panel, while it is the number of uses for the right panel.

Figure 7: Pattern of Price and Quantity Changes in a Sales Event



Note: The filled circles in black show observable posted price (top) and quantity purchased (bottom). The unfilled circles in red show unobservable effective price (top) and quantity consumed (bottom).

Figure 8: Simulated, Actual, and Inferred Paths of Price and Quantity



Note: The left-hand panel shows the simulated paths of price and quantity. The right-hand panel shows the actual paths of the price and the quantity purchased as well as those of the effective price and consumption for a cup noodle at a shop.

Figure 9: Histogram of the Degree of Stockpiling

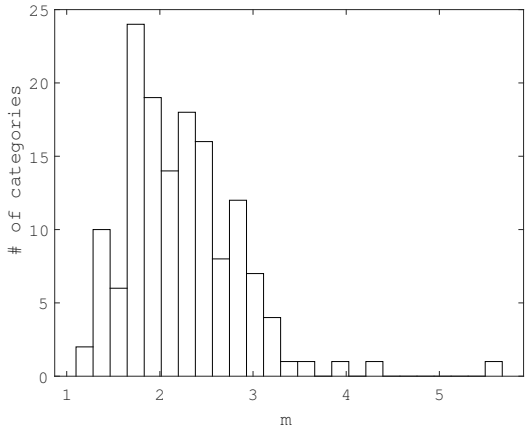
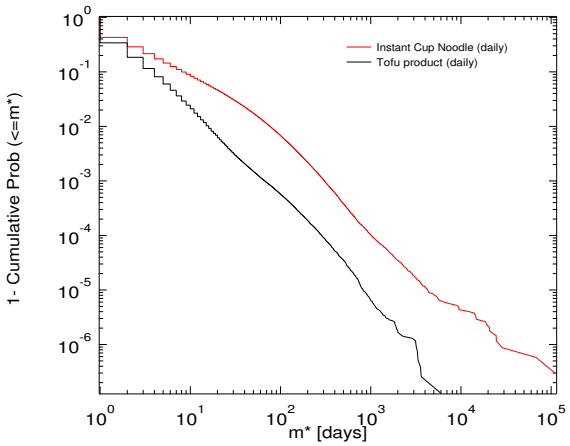
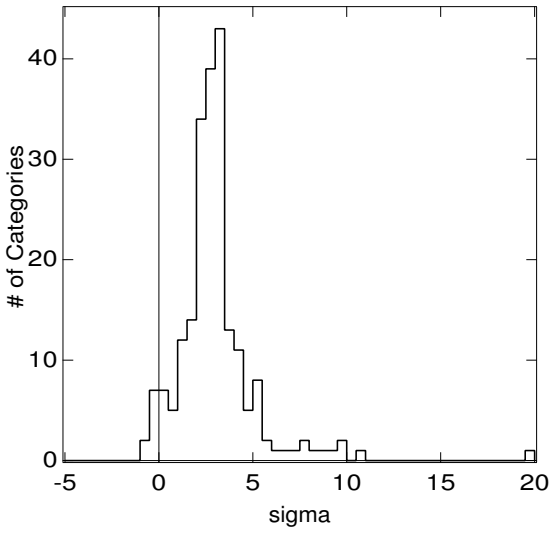


Figure 10: Degree of Stockpiling



Note: The figure shows the distribution of the inferred degree of stockpiling ($m < m^*$) for the cateogories of instant cup noodle and tofu product. In a cateogory, we have a sample for each product, shop, and sales event.

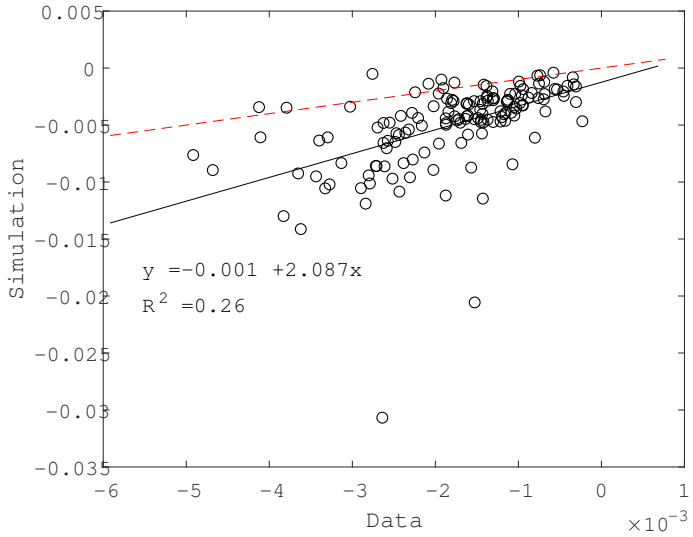
Figure 11: Elasticity of Substitution



Also for a naive estimate

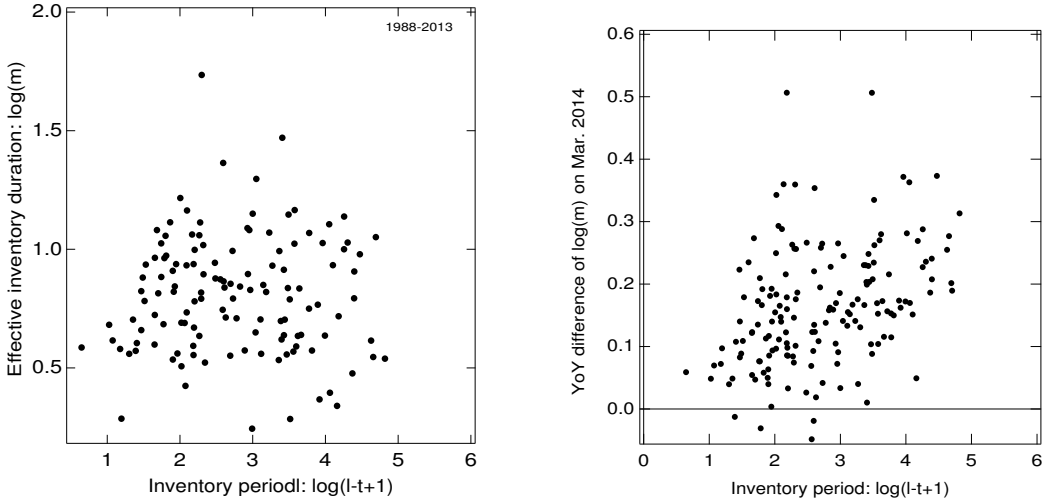
Note: The figure shows the histogram of the estimates of the elasticity of substitution for 3-digit product categories.

Figure 12: Inflation Rates Based on Törnqvist: Data and Simulation



Note: Each circle indicates a 3-digit product category. The inflation rates are daily average and based on the purchase-weighted Törnqvist index. The red dashed line represents the 45 degree line.

Figure 13: Relation between the Inventory Duration (Shokumap) and the Degree of Stockpiling (POS)



Note: Each dot indicates a 3-digit product category.

Figure 14: Relation between the Quantity of Purchase (Shokumap) and the Degree of Stockpiling (POS)

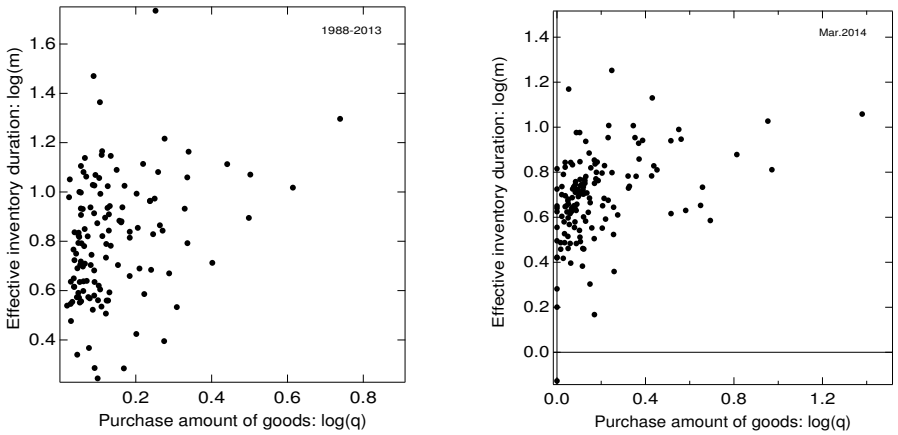
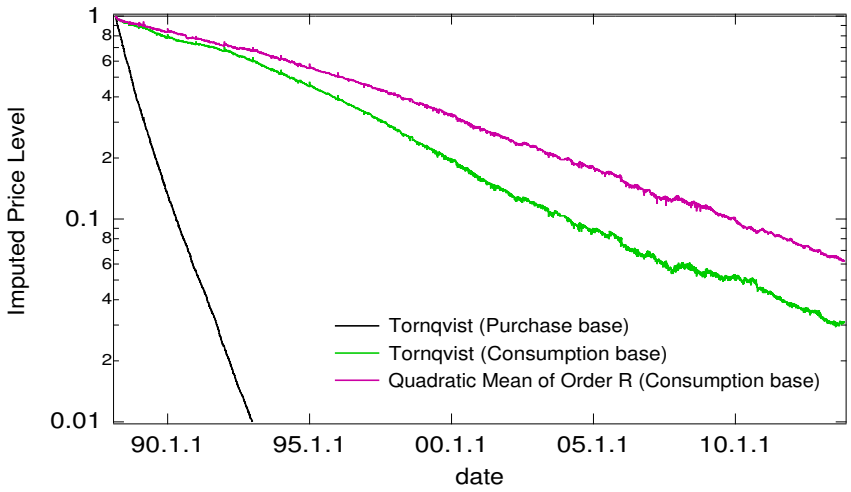
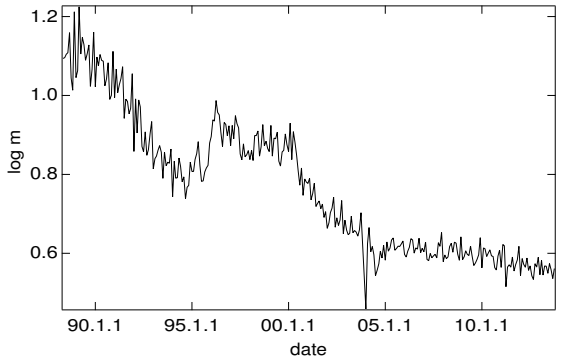


Figure 15: Price Indexes



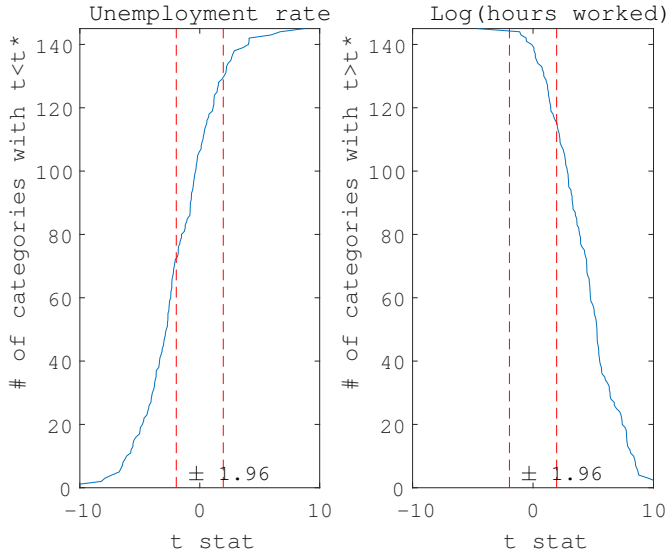
Note:

Figure 16: Aggregate Timeseries of the Degree of Stockpiling



Note:

Figure 17: Distribution for Business-Cycle Dependence of the Degree of Stockpiling



Note: We regress the degree of stockpiling for each 3-digit category level and obtain the t statistics for the coefficients on the unemployment rate and the hours worked. The vertical axis indicates the number of categories whose t statistics is smaller (larger for the right-hand panel) than t^*, while the horizontal axis represents t^*.