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

Kozo Ueda* Kota Watanabe† Tsutomu Watanabe‡

July 12, 2020

A Notes on Retailer Scanner Data

A.1 Aggregation of Variables

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

1. To construct the price indices and COLIs, we use the formula explained in the main text for products in each 3-digit product category. We assume that products are different when they are sold by different retailers even when the brand is identical, since the timing of sales and the degree of stockpiling differ across retailers. We then aggregate the price indices and COLIs at the 3-digit product category level using the sales weight of each product in the month that includes the day for which the aggregation is done.

2. To construct the aggregate variables except for the price indices and COLIs, we take the logarithm of variables (unless they are expressed by a rate of change or ratio) and then aggregate the values over products and retailers (and sales events), assigning equal weights. Such variables include $\log(m)$, $\log(1-\overline{q})$, $\log(1-\underline{q})$, and $\log((P_H - P_L)/P_H)$.

*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). All remaining errors are our own.

3. In both cases, we aggregate variables at the 3-digit category level only when the elasticity of substituion (σ) is greater then 1.0.

B Notes on the "***Shoku-map***" Household Scanner Data

Table 1 shows the basic statistics for the *Shoku-map* household scanner data. In the data, the definition of the number of products seems slightly ambiguous. For example, it seems that the number of products purchased is recorded as six, rather than one, when a household purchases a six-pack of beer. On the other hand, it seems that the number of chocolates purchased is recorded not as 24 but one when a household purchases a pack of 24 pieces of chocolate.

C Model Details

Proof of Lemmas 1 and 2

Define the Lagrangian as

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

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

$$0 = r_t - p_t + \psi_t, \tag{2}$$

$$C'(i_t) = \beta \mathrm{E}_t[r_{t+1}] - r_t + \mu_t. \tag{3}$$

The price to be paid by the consumer (consumer price), r_t, satisfies $0 < r_t \leq P_H$. The reason is that if r_t was strictly higher than P_H, consumers would purchase storable product k directly from product k manufacturers. Thus, P_H serves as the upper price limit.

Furthermore, r_t generally lies in the range of $[P_L, P_H]$. If $\mathrm{E}_t\,[r_{t+1}]$ was strictly lower than P_L, household producers would experience negative profits, since they purchase goods at P_L or P_H. Anticipating this possibility, household producers would not enter the market in the first place. However, a large *unexpected* shock may lead household producers to conduct fire sales to save inventory costs. Without such a large unexpected shock, when $p_t = P_L$, we have $r_t = r(I_{t-1}, P_L, b_t) = P_L$, because consumers can purchase storable goods directly from manufacturers.

Therefore, we are particularly interested in the consumer price at P_H, that is, $r(I_{t-1}, P_H, b_t)$. It is obvious that $r(I_{t-1}, P_H, b_t) \geq r(I_{t-1}, P_L, b_t) = P_L$. When $p_t = P_L$ and when $p_t = P_H$, equation (3) becomes

$$C'(i_L; I_{t-1}) = \beta \left\{(1 - \underline{q})r(I_t, P_H, b_t) + \underline{q}P_L\right\} - P_L + \mu_t, \tag{4}$$

$$C'(i_H; I_{t-1}) = \beta \left\{(1 - \overline{q})r(I_t, P_H, b_t) + \overline{q}P_L\right\} - r(I_{t-1}, P_H, b_t) + \mu_t, \tag{5}$$

respectively. Equation (4) indicates that, 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 household producers hold the same amount of inventories at $p_t = P_L$. At $p_t = P_H$, equation (5) 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 (5) 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 suggests 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 consumer price is expected to increase much, household producers hold some inventories at $p_t = P_H$. Otherwise, household producers do not hold inventories, that is, $i_H = 0$ at $p_t = P_H$. Inventory i_H is decreasing in $r(I_{t-1}, P_H, b_t)$. If $I_{t-1} = 0$, then $r(I_{t-1}, P_H, b_t) = P_H$, that is, consumers purchase directly from manufacturers.

Equations (4) and (5) 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 would be negative, while the left-hand sign 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 consumer demand is increasing in b_t according to the equation for the optimal quantity purchased,

$$c_t^k = \left(\frac{r_t^k / b_t^k}{r_{t'}^k / b_{t'}^k}\right)^{-\sigma} c_{t'}^k, \tag{6}$$

which is derived from consumers' cost minimization problem. At $p_t = P_L$, $r(I_{t-1}, P_L, b_t)$ is independent of b_t unless a large negative shock to 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 (2) suggests that $\psi_t > 0$, that is, $x_t = 0$, when $r_t < p_t$. In other words, household producers do not purchase 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 first part of Lemma 3, we examine the case of $p_t = p_{t+T+1} = P_H$. From the previous lemma, it follows that household producers do not purchase goods, i.e., $x_t = x_{t+T+1} = 0$. Because inventories at P_L, i_L, are independent of i_{t-1} and I_{t-1}, and inventories at P_H, i_H, are 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 consumers do not purchase goods from manufacturers when $r_t < P_H$, the quantity purchased by consumers is greater than or equal to that just after a sale.

For the second part of Lemma 3, consider a sales period from $t+1$ to $t+T$ $(T \geq 3)$. It is obvious that $I_t \leq I_{t+T-1}$. Because i_L is independent of i_{t-1} and I_{t-1}, the quantity purchased by household producers on the first day of a sale is greater than or equal to that on the final day of the sale. The quantity purchased by consumers is the same on the first and the last day of the sale, since consumers face the consumer price P_L. ■

Proof of Lemma 4

Proof. In the proof of Lemma 3, we showed that the inventory i_H at $p_t = P_H$ is decreasing over time. This means that, in aggregate, I_t is decreasing over time after a sale ends until I_t reaches zero. Since $r_t = r(I_{t-1}, p_t, b_t)$ is nonincreasing in I_{t-1}, r_t is nondecreasing over time. Equation (6) means that c_t is decreasing in r_t and therefore is nonincreasing over time. ■

Proof of Lemma 5

Proof. Note that the cost function, $\lambda_t = \mathcal{C}(r_t)$, is given by

$$\mathcal{C}(r_t) = \sum_{k \in K_t} r_t^k c_t^k = \left[\sum_{k \in K_t} \left(b_t^k\right)^{\sigma} \left(r_t^k\right)^{1-\sigma} U \right]^{1/(1-\sigma)}. \tag{7}$$

From equation (7), it follows that the cost function in each period can be 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 a constant. It can therefore be immediately seen that the change in the COLI from t to $t+T+1+T_H$, $\pi^{COLI} = \log \left\{ \Pi_{i=1}^{T+1+T_H} (\mathcal{C}(r_{t+i}) / \mathcal{C}(r_{t+i-1})) \right\}$, equals zero.

The total quantity purchased by household producers and consumers 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 $m \geq 1$, inventories at the end of a sale are not zero but positive, so $X_t > X_{t+T+1}$ and $X_{t+1} > X_{t+T}$. Thus, using $\sum_{k' \in K_0 \cap K_t} p_t^{k'} x_t^{k'} = 1$, we can show that the purchase-weighted Törnqvist index is given by

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

Similarly, we obtain

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

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

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

and

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

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 consumption price satisfies $P_L < r_{t+T+1} < P_H$ just after a sale 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 a sale) and $r_{t+j} = P_L$ for $j = 1, \cdots, T$ (i.e., during a sale).

We can show that the consumption-weighted Törnqvist index is given by

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

The consumption price drops from P_H to P_L in one period and takes more than one period to increase to P_H. If it takes two periods, π^{T*} is

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

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 exists a certain $r_H^*\subset(P_L,P_H)$ that makes $F'(r_H^*)$ equal 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 show here focuses on the case in which the consumption price decreases in one period and takes two periods to return to the original price, this logic also holds in cases where it takes longer than two periods for the price to return to the original price.

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 find that $G(r_H^*)$ is positive. Thus, $G(r_H)$ is positive. ■

Proof of Lemma 6

Proof. Suppose that the unit cost function is given by

$$\mathcal{C}(r_t)=\left[\sum_{i\in K}\sum_{k\in K}\alpha^{ik}\left(r_t^i\right)^{(1-\sigma)}\left(r_t^k\right)^{(1-\sigma)}\right]^{1/\{2(1-\sigma)\}} \tag{8}$$

where $\alpha^{ik}=\alpha^{ki}$. Using Shephard's Lemma and equation (8), we obtain

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

which yields

$$\frac{r_t^i c_t^i}{U_t\mathcal{C}(r_t)}=\frac{\left(r_t^i\right)^{1-\sigma}\sum_{k\in K}\alpha^{ik}\left(r_t^k\right)^{1-\sigma}}{\sum_{i\in K}\sum_{k\in K}\alpha^{ik}\left(r_t^i\right)^{(1-\sigma)}\left(r_t^k\right)^{(1-\sigma)}}. \tag{9}$$

Noting that

$$\begin{aligned}
s_t^i &\equiv \frac{r_t^i c_t^i}{\sum_{k\in K} r_t^k c_t^k}\\
&= \frac{r_t^i c_t^i}{U_t\mathcal{C}(r_t)}\\
&= \frac{\left(r_t^i\right)^{1-\sigma}\sum_{k\in K}\alpha^{ik}\left(r_t^k\right)^{1-\sigma}}{\sum_{i\in K}\sum_{k\in K}\alpha^{ik}\left(r_t^i\right)^{(1-\sigma)}\left(r_t^k\right)^{(1-\sigma)}},
\end{aligned} \tag{10}$$

we obtain

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

■

Equilibrium When Inventories are Cleared in Just One Period After a Sale

Household producers' firm 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 yields

$$V(i_{t-1} = 0, p_t = P_L; I_{t-1}) = 0, \tag{11}$$

while $V(i_{t-1} = 0, p_t = P_H; I_{t-1}) < 0$. Using this condition, we can show that r_H is a function of I_{t-1}, that is, $r_H(I_{t-1})$.

If household producers 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 can be 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 (4) can be rewritten as

$$C'(i_L) + (1 - \beta \underline{q}) P_L = \beta(1-\underline{q}) r_H(I_L), \tag{12}$$

we obtain

$$C'(i_L) i_L - C(i_L) = \frac{\beta(1-\underline{q})}{1-\beta(1-\overline{q})} C(0). \tag{13}$$

Equations (12) and (13) 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 $p_t = P_L$, manufacturers supply goods to consumers with no additional cost until the goods market is cleared, because consumption price r_t is equal to P_L. In the case of $p_t = p_{t-1} = P_H$, we have $r_t = P_H$. Consumers again purchase goods directly from manufacturers until the goods market is cleared. In the case of $p_t = P_H$ and $p_{t-1} = P_L$, consumption price r_t equals r_H and the quantity of goods supplied (which equals the quantity consumed) is predetermined because inventories are predetermined and there are no additional purchases (i.e., $x_t = 0$). Thus, if demand in the goods market is greater (smaller) than supply, r_H increases (decreases) until the market is cleared, which makes household producers' firm value deviate from its expected value. In other words, the expected consumption price at $t+1$, r_{t+1}, conditional on $p_t = P_H$ and $p_{t-1} = P_L$, equals r_H, but the realized consumption price may differ from the expected value. Also, it should be pointed out that r_H is independent of factors on the household side, that is, b_t^k. Such factors influence the demand for storable product k but do not influence the consumption price r_H, since household producers freely enter the market and change the aggregate supply of goods.

D Detailed Explanation of Our Approach

D.1 Proof of m_{cont}

Assume time is continuous. If r_t increases linearly in 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 elapsed after a sale, because $r_H(0) = P_L$ and $r_H(m_{cont}) = P_H$. The corresponding consumption equals $(r_H/P_L)^{-\sigma} c_L^*$. Denote the initial inventories outstanding just after a sale ends by I_L. At $m = m_{cont}$, all inventories are used up, which is given by

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

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

D.2 Conditions under Which Our Approach Holds

When $p_t = P_H$, household producers optimize inventories to satisfy

$$C'(i_H; I_{t-1}) = \beta \left\{ (1 - \overline{q}) r_H(I_t) + \overline{q} P_L \right\} - r_H(I_{t-1}) + \mu_t. \tag{14}$$

Positive inventories yield $\mu_t = 0$. Then, $r_H(I_t) - r_H(I_{t-1})$ in equation (14) becomes a positive constant if

$$C'(i_H; I_{t-1}) + \left\{ 1 - \beta(1 - \overline{q}) \right\} r_H(I_{t-1}) \tag{15}$$

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

Equation (15) suggests that $C(\cdot)$ should satisfy the following conditions. Note that if $r_H(I_t) - r_H(I_{t-1})$ is a positive constant, the consumption price and inventories at time x ($0 \le x \le m_{cont}$) equal

$$r_H(x) = \left\{ \frac{x}{m_{cont}} \frac{P_H - P_L}{P_L} + 1 \right\} P_L,$$

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

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

respectively. Thus, inserting $i_H = \kappa I_t$ into equation (15), we can show that the following term,

$$C'\left(\left\{ 1 - \frac{1 - \left(\frac{x}{m_{cont}} \frac{P_H - P_L}{P_L} + 1 \right)^{-\sigma+1}}{1 - (P_H/P_L)^{-\sigma+1}} \right\} \kappa I_L \right) + \left\{ 1 - \beta(1 - \overline{q}) \right\} \left\{ \frac{x}{m_{cont}} \frac{P_H - P_L}{P_L} + 1 \right\} P_L, \tag{16}$$

needs to be independent of x 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 the following condition for the inventory cost function:

$$C'(x') = C_0 - \{1 - \beta(1 - \overline{q})\} \left\{ 1 - \left(1 - \frac{x'}{\kappa I_L} \right) \left(1 - \left(\frac{P_H}{P_L} \right)^{-\sigma+1} \right) \right\}^{\frac{1}{-\sigma+1}} P_L, \tag{17}$$

where $C_0 > \beta \overline{q} P_L$.[1] It can be easily shown that $C''(\cdot) > 0$ provided $\sigma > 1$.

D.3 Details of How We Apply Our Approach to the POS Data

Using the POS data, we calculate $r_H(I_{t-1})$, m, X_L^*, and I_L for each sales event for each product at each retailer, while we calculate σ for each 3-digit product category.

We take the following three steps. First, for each product and retailer, we identify all sales events using the sales filter.[2] We define the first and the last day of a sales event as $t + 1$ and $t + T$, respectively. We record the price and the quantity purchased just before a sale (i.e., at t) as r_H and c_H, respectively. Note that, according to the model, c_H should equal consumption unless the effect of stockpiling remains and c_H is zero. In that case, we use the previous values of r_H and c_H.[3]

Second, if a sale lasts more than one day ($T > 1$), we record the average price and the quantity purchased in the first half of a sale (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 of the sale (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 this case, there is no need for additional stockpiling in the second half of the sale, so X_L^2 should equal consumption, which we denote by c_L. However, the POS data often show the opposite,

[1] The inventories outstanding just after a sale ends, I_L, are endogenous and determined from the optimization equation when $p_t = P_L$:

$$C'(\kappa I_L) = \beta \left\{ (1 - \underline{q}) r_H(x = 1) + \underline{q} P_L \right\} - P_L.$$

[2] Before identifying sales events, we remove outliers by resetting the price and quantity purchased to zero for product i on day t if the absolute log difference between p_{it} and the median of p_{it} over the observation period is larger than log(70/10). Additionally, if the number of days with $p_{it} > 0$ is smaller than seven, we remove this product from our study.

[3] In order to avoid a situation such that the first observed c_H is zero, we search for the earliest sales event that satisfies $c_H > 0$ and then go back in time as far as the regular price r_H continues to be observed. We define the first day when the regular price is listed as the starting point of the time series for estimating consumption paths.

$X_L^1 < X_L^2$, which makes sense if the duration of the sale is known *ex ante*. In this case, there is no need for stockpiling in the first half of the sale, so that X_L^1 should equal c_L.

In the third step, we calculate the elasticity of substitution, σ. The variables we obtained in the first two steps, $(r_H,\, r_L)$ and $(c_H,\, c_L)$, correspond to the consumption prices and the quantities consumed. Thus, equation (6) should hold for these variables when the true σ is used. Furthermore, the log ratio of the quantity consumed during a sale to that when the product is sold at the regular price divided by the log ratio of the sales price to the regular price, $\Gamma \equiv -\log\left(c_L/c_H\right)/\log\left(r_L/r_H\right)$, should equal σ on average. Thus, we calculate the unweighted average of Γ across sales events, products, and retailers for each 3-digit product category, which we define as σ.

Fourth, we calculate consumption, the consumption price, the degree of stockpiling, and inventories for each sales event, each product, and each retailer. If a sale lasts more than one day ($T > 1$), we set $c_L^* = \left(r_L/r_H\right)^{-\sigma} c_H$. If a sale is only one day long, we set $r_L = p_t$ and $c_L^* = \min\left[0.01, \left(r_L/r_H\right)^{-\sigma} c_H\right]$. Inventories at the end of a sale, I_L, equal the cumulative quantity of purchases during the sale minus the cumulative quantity of consumption during the sale, that is, $\max\left[0, \Sigma_{j=1}^{T} X_{t+j} - T c_L^*\right]$, where we replace I_L with zero if the calculated amount of inventories is negative. Once we have obtained c_L^* and I_L, we can calculate m. Note that since the unit of time in our analysis is discrete (i.e., we use daily observations), we search for the maximum natural number m starting from the above continuous-time version of m_{cont}, so that inventories at $t+T+m$ are positive, while those at $t+T+m+1$ are negative. Once we have obtained m, we can calculate the path of consumption prices $r_s = r_H(I_{s-1})$ and consumption c_s as $r_L + ((t'-t-T)/m)(r_H - r_L)$ and $\left(r_s/r_L\right)^{-\sigma} c_L^*$, respectively, for $t+T+1 \leq t' \leq t+T+m$.[4]

[4] It is possible that the next sale begins at $t^{'}$ before $t+T+m$. In this case, we use r_s and c_s until $t^{'}-1$ and recalculate them for $t' \geq t^{'}$ assuming that inventories are zero ($I_{t^{'}-1} = 0$) when the new sale starts. We use the same c_H because information on the quantity purchased just before the sale is not updated.

E Simulation Results and Comparison with the POS Data

E.1 Simulation

Table 2 shows that the size of the chain drift increases as the degree of stockpiling, m, increases, the size of the sale discount increases (P_L/P_H decreases), and the probability that a product will go on sale on the following day given that it is not currently on sale, $\overline{q}$, increases. On the other hand, interestingly, an increase in the probability that a product will continue to be on sale on the following day given that it is currently on sale, $\underline{q}$, decreases the size of the chain drift in the purchase-weighted Laspeyres and Paasche indices, but not in the purchase-weighted Törnqvist index. The like reason is that the expectation of longer sales reduces the incentive to stockpile.

E.2 Comparison with the POS Data

To validate our approach, we examine whether the size of the simulated chain drift is comparable to that of the actual chain drift. To do so, 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 (denoted by π_j). The first and last 12-month periods of the data are omitted from the calculation because identifying sales events is difficult when data are censored.

We regress the average inflation rate π_j based on the POS data on $\log(m_j)$, $\overline{q_j}$, $\underline{q_j}$, and $\log(P_L/P_H)_j$, which are also obtained from the POS data. Table 3 shows the estimation results. 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. These results are all consistent with the simulation results shown in Table 2; that is, we find that a high m, a high $\overline{q}$, and a low P_L/P_H all lead to greater deflation, while the effect of $\underline{q}$ is negligible.

F Other Approaches

We examine how much the size of the chain drift changes when we use different assumptions with regard to the consumption price. Instead of assuming a linear consumption-price increase, we use two types of alternative approaches, allowing for some convexity or concavity in the consumption-price increase.

F.1 Linear Consumption Decrease

The first alternative approach is to assume a linear consumption decrease after a sale ends until inventories decrease to zero. Assume time is continuous. Denote consumption before and during a sale by c_H and c_L^*, respectively (note that $c_H < c_L^*$). If consumption c_t decreases linearly in t after a sale, we can write c_t as $c(x) = \left\{ \frac{x}{m_{cont}} \frac{c_H - c_L^*}{c_L^*} + 1 \right\} c_L^*$, where x represents the time elapsed after a sale ($0 \leq x \leq m_{cont}$), because $c(0) = c_L^*$ and $c(m_{cont}) = c_H^*$. At $m = m_{cont}$, all inventories are used up, which is given by

$$
\begin{aligned}
I_L &= \int_0^m \left\{ \frac{x}{m_{cont}} \frac{c_H - c_L^*}{c_L^*} + 1 \right\} c_L^* dx \\
\therefore m_{cont} &= \frac{2}{c_L^* + c_H} I_L. \qquad (18)
\end{aligned}
$$

The corresponding consumption price $r(x)$ can be written as

$$
r(x) = \left\{ \frac{x}{m_{cont}} \frac{c_H - c_L^*}{c_L^*} + 1 \right\}^{-1/\sigma} P_L. \qquad (19)
$$

It can be easily shown that $r'(x) > 0$ and $r''(x) < 0$. Thus, the consumption price increases in x, and the speed of the consumption-price increase decreases in x (i.e., the consumption-price increase is concave).

F.2 Convex or Concave Consumption-price Increase

As the second alternative approach, we add or subtract a particular integer ε from m that is derived from our baseline approach. Denoting the new m by $m' = m + \varepsilon$, we calculate parameters γ_0 and γ_1 such that it satisfies $r(x) = \left(P_L^{1/\gamma_0} + \gamma_1 x \right)^{\gamma_0}$, $r(m') = P_H$, and $I_L = \int_0^{m'} (r(x)/P_L)^{-\sigma} c_L^* dx$. In other words, γ_0 and γ_1 satisfy

$$
\begin{aligned}
P_H &= \left(P_L^{1/\gamma_0} + \gamma_1 m' \right)^{\gamma_0}, \\
I_L &= \left[\left(\frac{P_L^{1/\gamma_0} + \gamma_1 m'}{P_L} \right)^{-\gamma_0\sigma+1} - \left(\frac{P_L^{1/\gamma_0}}{P_L} \right)^{-\gamma_0\sigma+1} \right] c_L^* \frac{P_L}{(-\gamma_0\sigma + 1)\gamma_1}. \qquad (20)
\end{aligned}
$$

Note that when $\varepsilon = 0$, γ_0 equals 1. If $\gamma_0 > 1$ and $\gamma_1 > 0$, the speed of the consumption-price increase increases in x (i.e., the consumption-price increase is convex). If $0 < \gamma_0 < 1$ and $\gamma_1 > 0$, the speed of the consumption-price increase decreases in x (i.e., the consumption-price increase is concave).

F.3 Simulation

We calculate the price indices based on the COLI, the chained order r superlative, the chained consumption-weighted Törnqvist, the chained purchase-weighted Törnqvist, the chained purchase-weighted Laspeyres, and the chained purchase-weighted Paasche, using the method explained in Section 4.2 in the main text. The benchmark value for the degree of stockpiling is $m = 5$.

Before simulating the price indices using the alternative approaches, we calculate the mean values of c_H, c_L^*, and I_L from the simulation results of $T = 365$ days times $N = 100$ in the benchmark case to use. In particular, using the value of I_L is important, because we are interested in examining how much the size of the chain drift changes when we assume different paths of consumption and of the consumption price after a sale ends, given a certain amount of inventories during a sale.

When we use the first alternative approach, we calculate m_{cont} from $2I_L/(c_L^* + c_H)$ and the maximum integer of m to satisfy $m \leq m_{cont}$. We then calculate the path of the consumption price from equation (19) (while $0 \leq x \leq m$), followed by the path of consumption so that it is consistent with the demand function.[5]

When we use the second alternative approach, we assume $m' = m + \varepsilon$, where ε takes -2, -1, 1, or 2. Solving equation (20), we respectively obtain values of 0.09, 0.46, 1.44, and 1.80 for γ_0. This suggests that when inventories are cleared in a shorter time than $m = 5$, we have $0 < \gamma_0 < 1$; i.e., the consumption-price increase is concave. When inventories are cleared in a longer time than $m = 5$, we have $\gamma_0 > 1$; i.e., the consumption-price increase is convex.

[5] We do this, rather than calculating the path of consumption followed by the consumption price, since consumption c_t^k of product k should depend on the consumption prices of other products.

Table 1: Basic Statistics of the *Shoku-map* Household Scanner Data

Variable	Mean	S.D.	Min	Max
# of households per month	358	78	105	450
# of products purchased per month	44,260	11,586	9,126	64,660
# of products purchased per month and per household	123	12	85	152
# of months for which a household answered	24	13	1	47
Age of the wife in the household	44	11	21	72

Note: # of households per month is defined as the number of households that made purchases in each month. # of products purchased per month is defined as the number of records in each month. # of months for which a household answered is the number of months for which a household made purchases.

Table 2: Simulation of the Chain Drift

	COLI	Order r superlative	Törnqvist (C)	Törnqvist (X)	Laspeyres (X)	Paasche (X)
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: The table shows the means of the price levels after 365 days, where the initial price level is set to one (so that a value of one indicates no change). Standard deviations in parentheses.

Table 3: 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 * denote significance at the 1%, 5%, and 10% levels, respectively.