Estimating Daily Inflation Using Scanner Data: A Progress Report

Kota Watanabe
Meiji University
& Univ of Tokyo

Tsutomu Watanabe
University of Tokyo

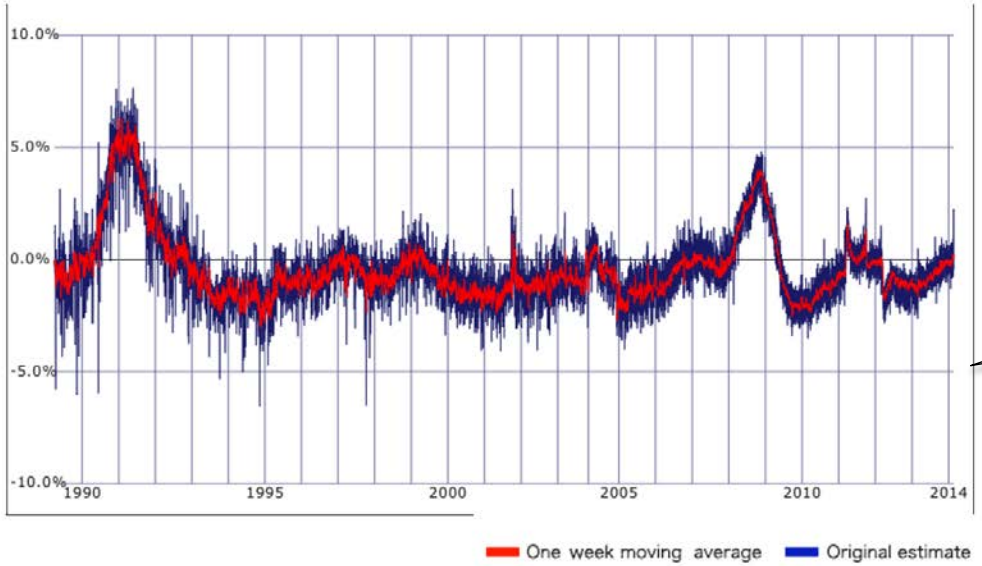
May 27, 2014

UTokyo Daily Price Project

The Daily Index for the most recent period has been updated on 2014-02-24

Daily Index | Monthly Index | FAQ

Daily Index > Nationwide



Original estimate:	-0.23%
One week moving average:	0.11%
Year-to-year inflation rate for 2014-02-21	

Daily Index

- Prices are collected every day from **300 supermarkets** sampled from all over Japan.
- The number of products is about 300K.
- Updated every day only with a **two day lag**.
- Covers **17 percent** of the official CPI.
- Historical daily data is available over the last 25 years.

The current and historical data are downloadable. **Free!**

http://www.cmdlab.co.jp/price_u-tokyo/dailys_e

How is the UTokyo Daily Price Index calculated?

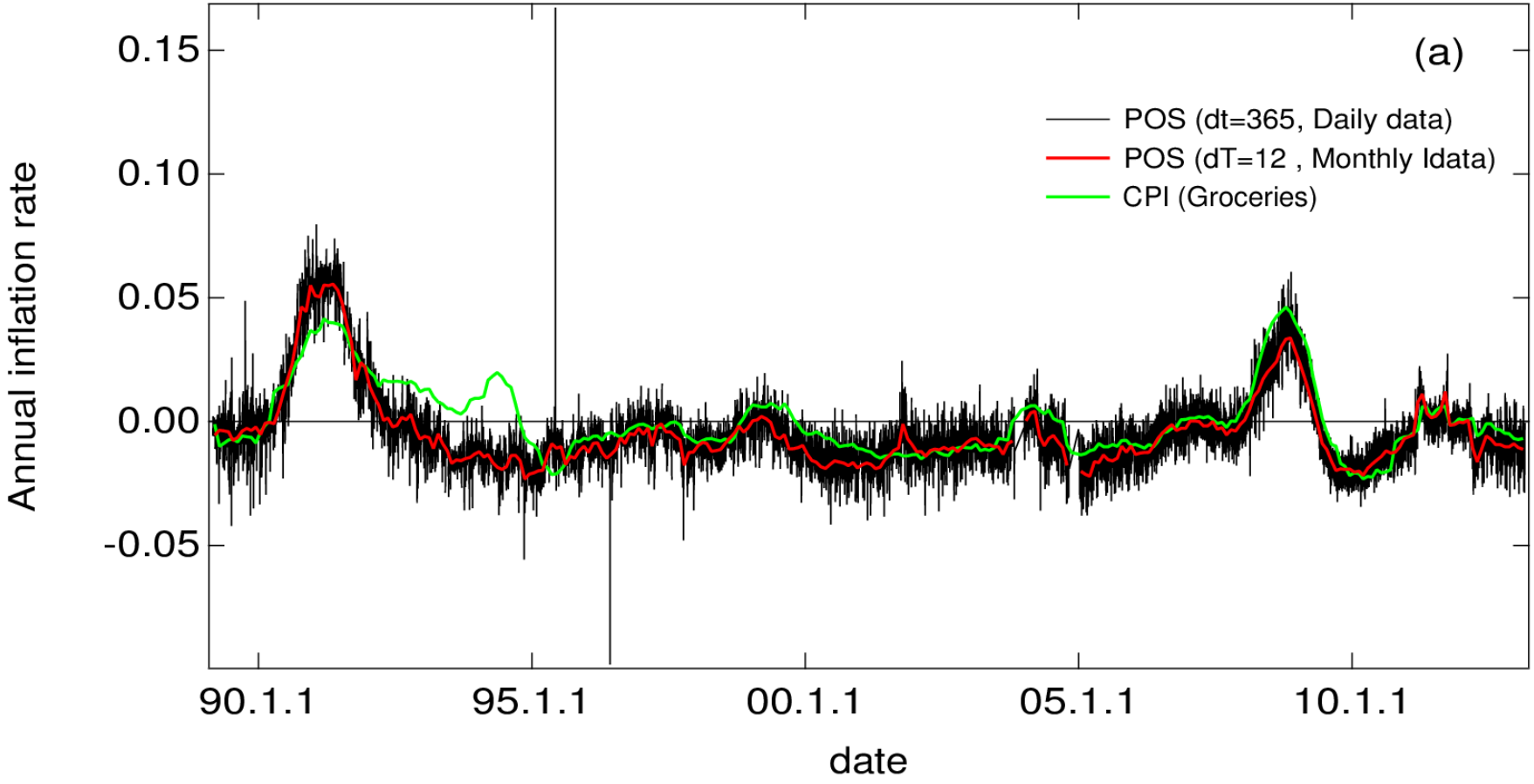
- The UTokyo Daily Price Index is **a daily version of the Törnqvist index**, which is known as one of the **superlative price indexes**.
  - CPI Manual released by ILO: "Many different kinds of mathematical formulae have been proposed over the past two centuries. While there may be no single formula that would be preferred in all circumstances, most economists and compilers of CPIs seem to be agreed that, in principle, the index formula should belong to a small class of indices called superlative indices." (Consumer price index manual: Theory and practice, 2004, p.2)

- **Price relatives**: The daily inflation rate is calculated as the weighted geometric mean of price relatives across products, which are defined as the price ratios **between today and the same day of the previous year**.

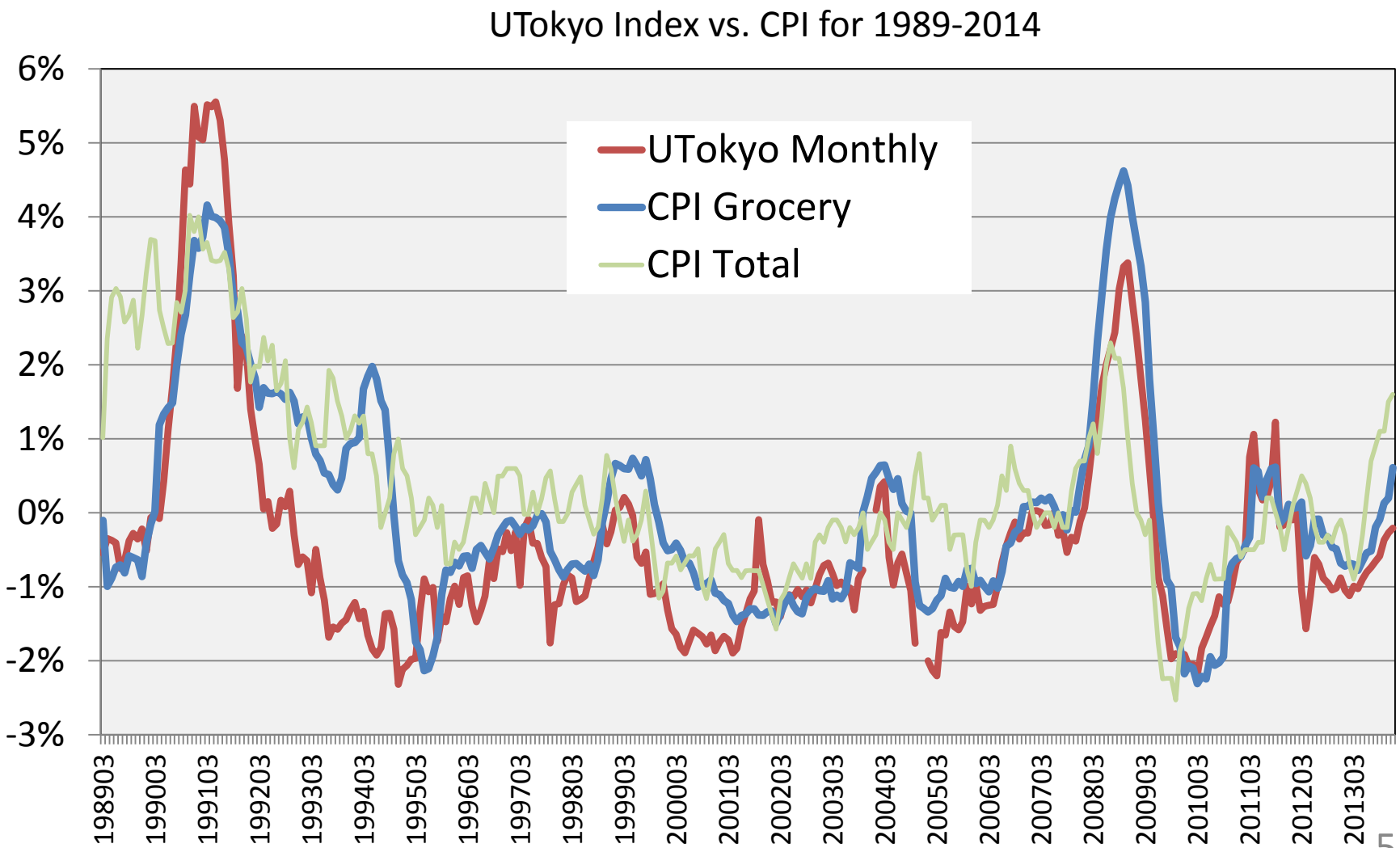
- **Weights**: The weight for a product is given by the average of the sales shares of the product today and the sales share of the same product on the same day of the previous year.

$$\pi_{t,t-dt} = \sum_i \frac{s_{i,t} + s_{i,t-dt}}{2} \ln\left(\frac{p_{i,t}}{p_{i,t-dt}}\right)$$

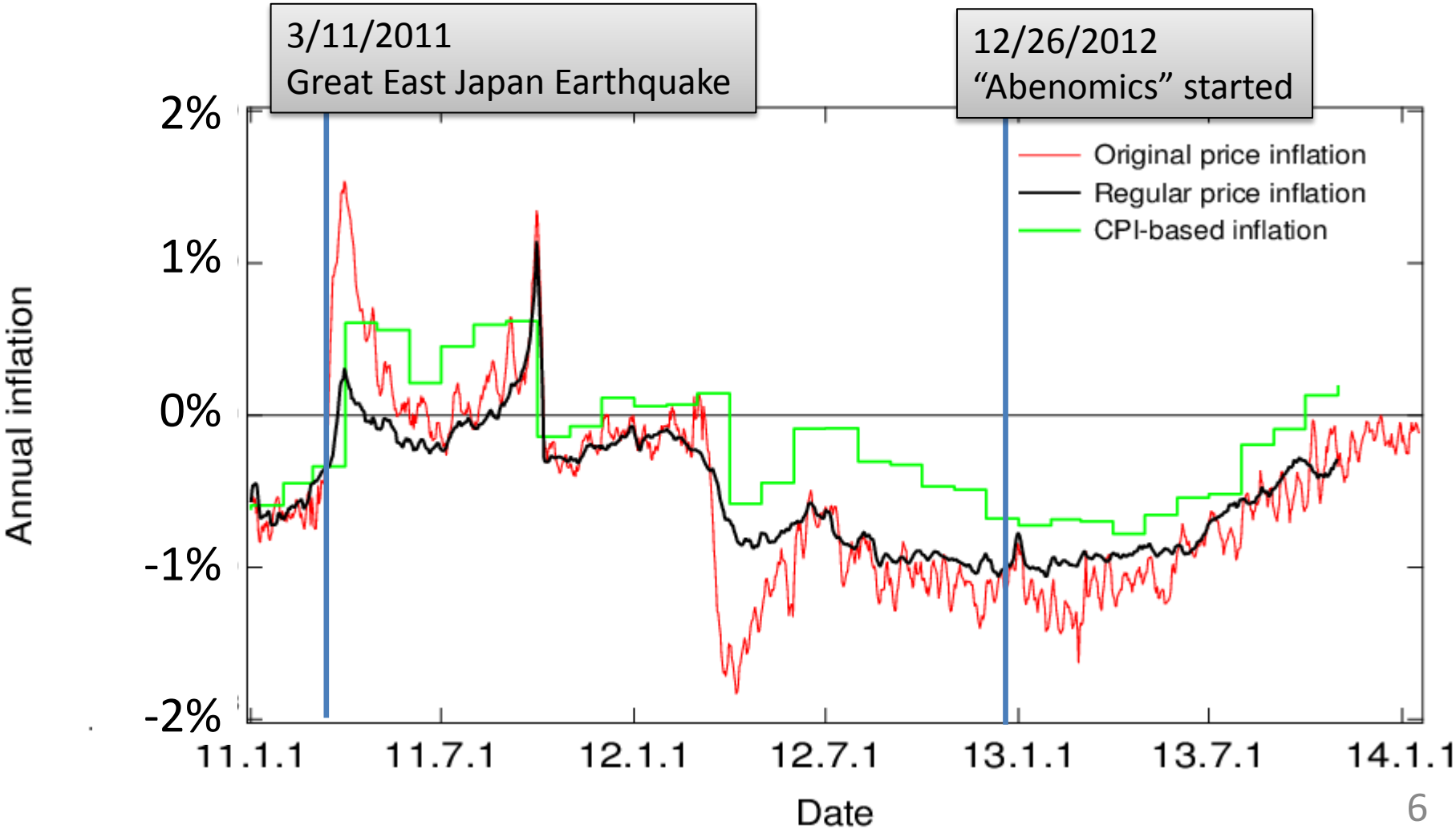
Result #1: Daily inflation is highly volatile



Result #2: Tornqvist inflation is lower than CPI inflation by half a percent



Result #3: High frequency inflation measure is useful to detect the impacts of various shocks, including policy shocks, on inflation



- Daily inflation by item
- Daily inflation by region
- Core inflation
- **Chained daily Tornqvist index**
- Daily sales index
- Decomposition of daily inflation into regular and sale price changes

Chaining

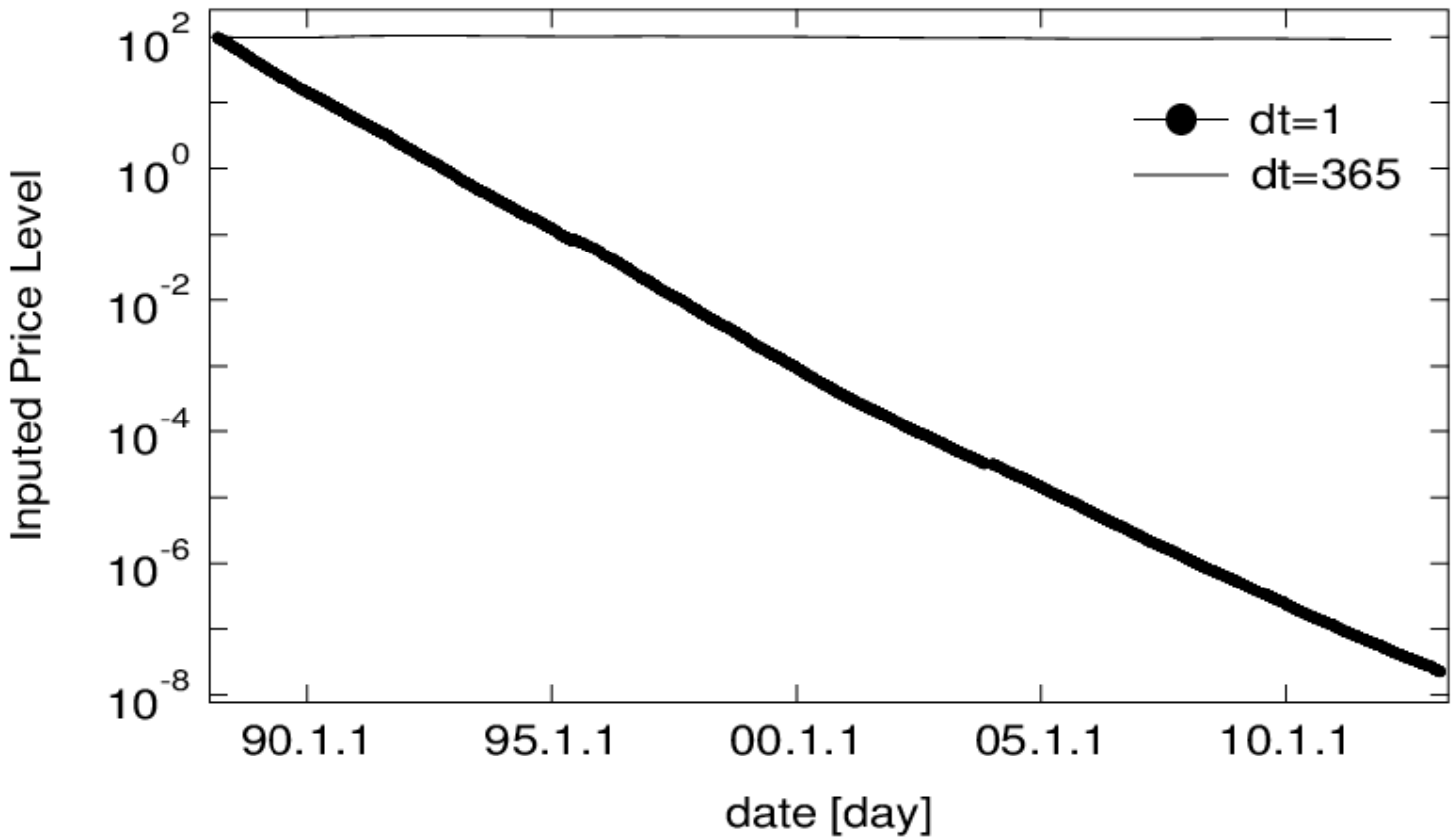
- The current version of UTokyo Daily Price Index is estimated by setting ***dt*=365**, i.e., by comparing the price of a product on a particular day and the price of the product on the same day in the previous year.

- The set of products used in this calculation consists of those products available on both that particular day and the same day a year earlier. Products that disappeared sometime between the two days are not included in the calculation. Similarly, products that were newly born somewhere between the two days are also not included.

- However, product turnover is very high in scanner data. The entry and exit rates in our dataset are 36 percent and 35 percent, respectively.

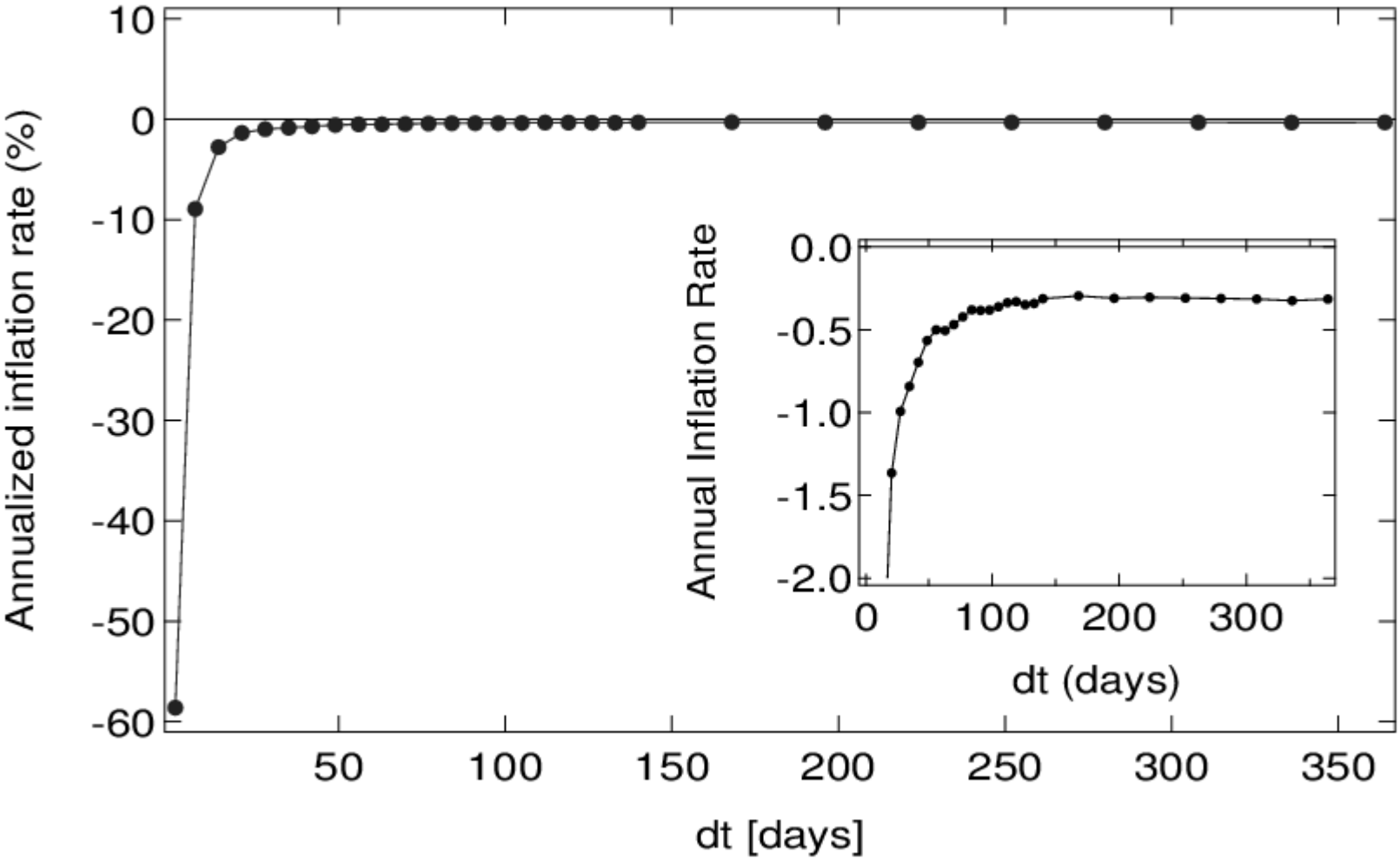
- We would like to maximize the number of matches by setting ***dt* = 1**, where the prices of products on a particular day and the day immediately before are compared.

Chained daily Tornqvist index

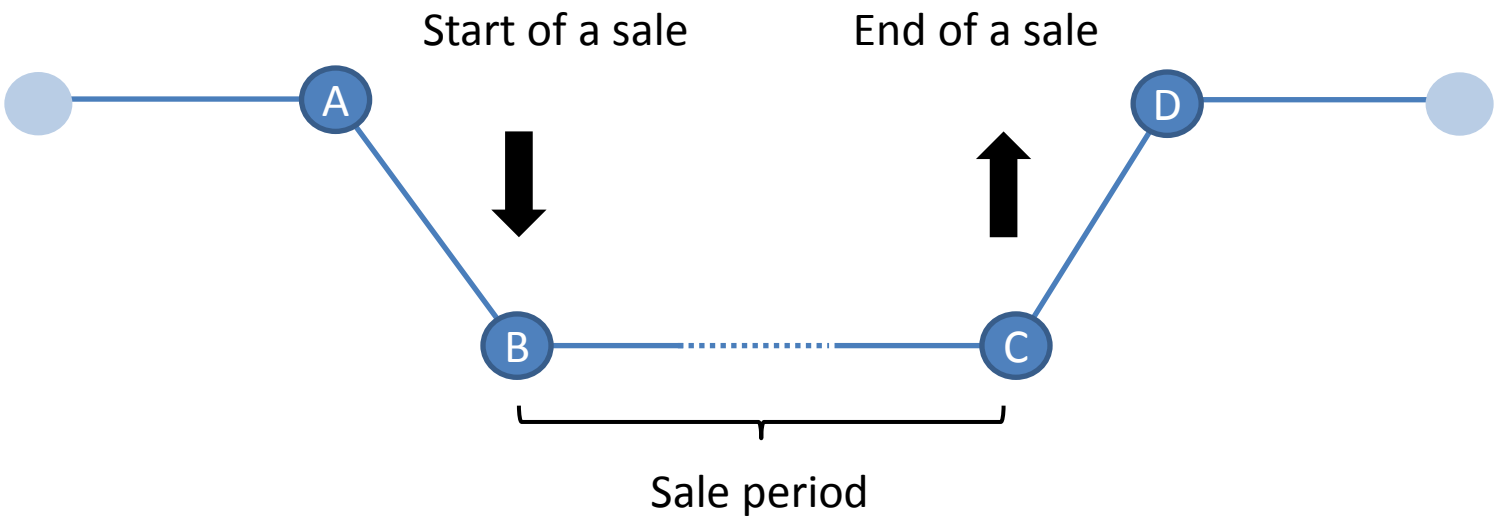


The chained price index falls to 10^{-10} of the base value over the 25 year sample period, which is equivalent to an annual deflation rate of 60 percent.

Time interval (*dt*) and the measured inflation rate

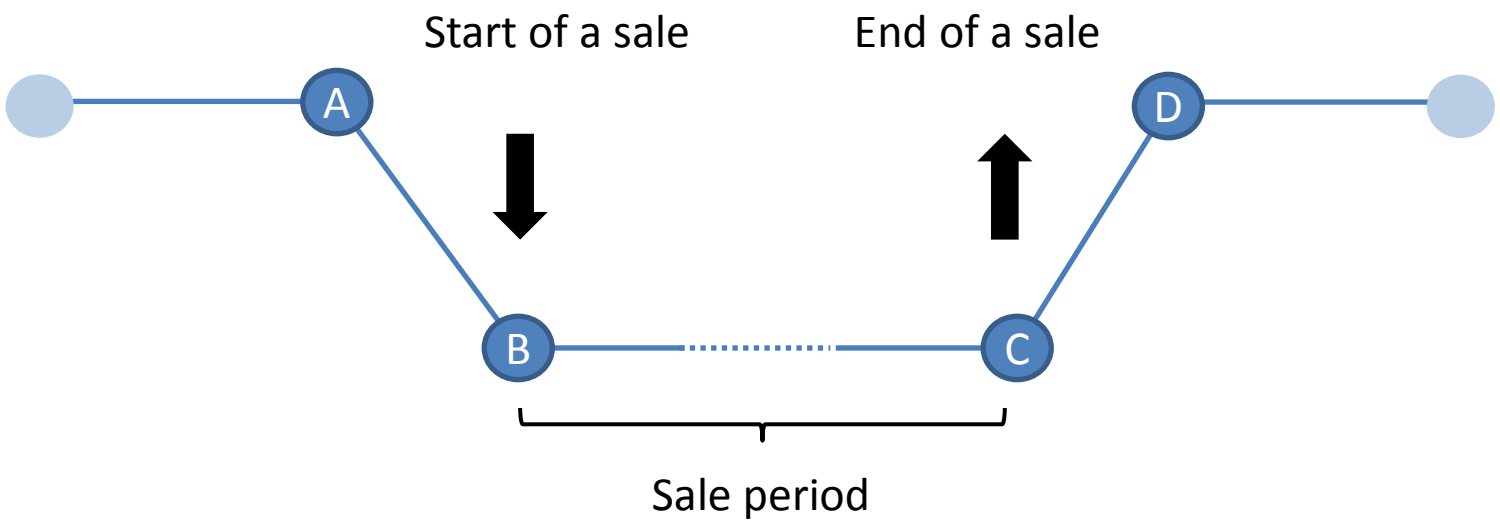


Why do we have chain drift?



A: Expenditure share on the day **before** a sale period
B: Expenditure share on the **first** day of a sale period
C: Expenditure share on the **final** day of a sale period
D: Expenditure share on the day **after** a sale period

- If goods are *not* storable, intertemporal utility maximization implies that $A=D$ and $B=C$, so that $A+B=C+D$. This symmetry implies that the Tornqvist weight associated with the price decline at the start of a sale, $(A+B)/2$, is the same as the Tornqvist weight associated with the price increase at the end of a sale, $(C+D)/2$. Therefore, no chain drift arises.



- However, this no longer holds if goods are storable.
  - Triplett (2003)
- Chain drift arises if $A+B \neq C+D$.

If $A>D$ & $B=C$ => Chained Tornqvist index has a downward drift

- Ivancic, Diewert, and Fox (2011); De Haan and van der Grient (2011)

If $A<D$ & $B=C$ => Upward drift

If $A=D$ & $B>C$ => Downward drift

If $A=D$ & $B<C$ => Upward drift

- Feenstra and Shapiro (2003)

$$
\begin{aligned}
\tilde{\pi}_t \;=\; & \frac{1}{2}\underbrace{\frac{\sum_{(i,s)\in R_t} e_{i,s,t}}{\sum_s \sum_i e_{i,s,t}}}_{\text{Expenditure share}} \underbrace{\sum_{(i,s)\in R_t} \frac{e_{i,s,t}}{\sum_{(i,s)\in R_t} e_{i,s,t}} \ln \frac{p^R_{i,s,t}}{p^R_{i,s,t-1}}}_{\text{Size of reg price change}} \\
+ \; & \frac{1}{2}\underbrace{\frac{n^D_t}{n_t}}_{f^D_t} \underbrace{\frac{\sum_{(i,s)\in D_t} e_{i,s,t}/n^D_t}{\sum_s \sum_i e_{i,s,t}/n_t}}_{x^D_t} \underbrace{\sum_{(i,s)\in D_t} \frac{e_{i,s,t}}{\sum_{(i,s)\in D_t} e_{i,s,t}} \left(\ln \frac{p_{i,s,t}}{p_{i,s,t-1}} - \ln \frac{p^R_{i,s,t}}{p^R_{i,s,t-1}} \right)}_{u^D_t} \\
+ \; & \frac{1}{2}\underbrace{\frac{n^D_t}{n_t}}_{f^D_t} \underbrace{\frac{\sum_{(i,s)\in D_t} e_{i,s,t-1}/n^D_t}{\sum_s \sum_i e_{i,s,t-1}/n_t}}_{y^D_t} \underbrace{\sum_{(i,s)\in D_t} \frac{e_{i,s,t-1}}{\sum_{(i,s)\in D_t} e_{i,s,t-1}} \left(\ln \frac{p_{i,s,t}}{p_{i,s,t-1}} - \ln \frac{p^R_{i,s,t}}{p^R_{i,s,t-1}} \right)}_{v^D_t} \\
+ \; & \frac{1}{2}\underbrace{\frac{n^U_t}{n_t}}_{f^U_t} \underbrace{\frac{\sum_{(i,s)\in U_t} e_{i,s,t}/n^U_t}{\sum_s \sum_i e_{i,s,t}/n_t}}_{x^U_t} \underbrace{\sum_{(i,s)\in U_t} \frac{e_{i,s,t}}{\sum_{(i,s)\in U_t} e_{i,s,t}} \ln \left(\ln \frac{p_{i,s,t}}{p_{i,s,t-1}} - \ln \frac{p^R_{i,s,t}}{p^R_{i,s,t-1}} \right)}_{u^U_t} \\
+ \; & \frac{1}{2}\underbrace{\frac{n^U_t}{n_t}}_{f^U_t} \underbrace{\frac{\sum_{(i,s)\in U_t} e_{i,s,t-1}/n^U_t}{\sum_s \sum_i e_{i,s,t-1}/n_t}}_{y^U_t} \underbrace{\sum_{(i,s)\in U_t} \frac{e_{i,s,t-1}}{\sum_{(i,s)\in U_t} e_{i,s,t-1}} \left(\ln \frac{p_{i,s,t}}{p_{i,s,t-1}} - \ln \frac{p^R_{i,s,t}}{p^R_{i,s,t-1}} \right)}_{v^U_t}
\end{aligned}
$$

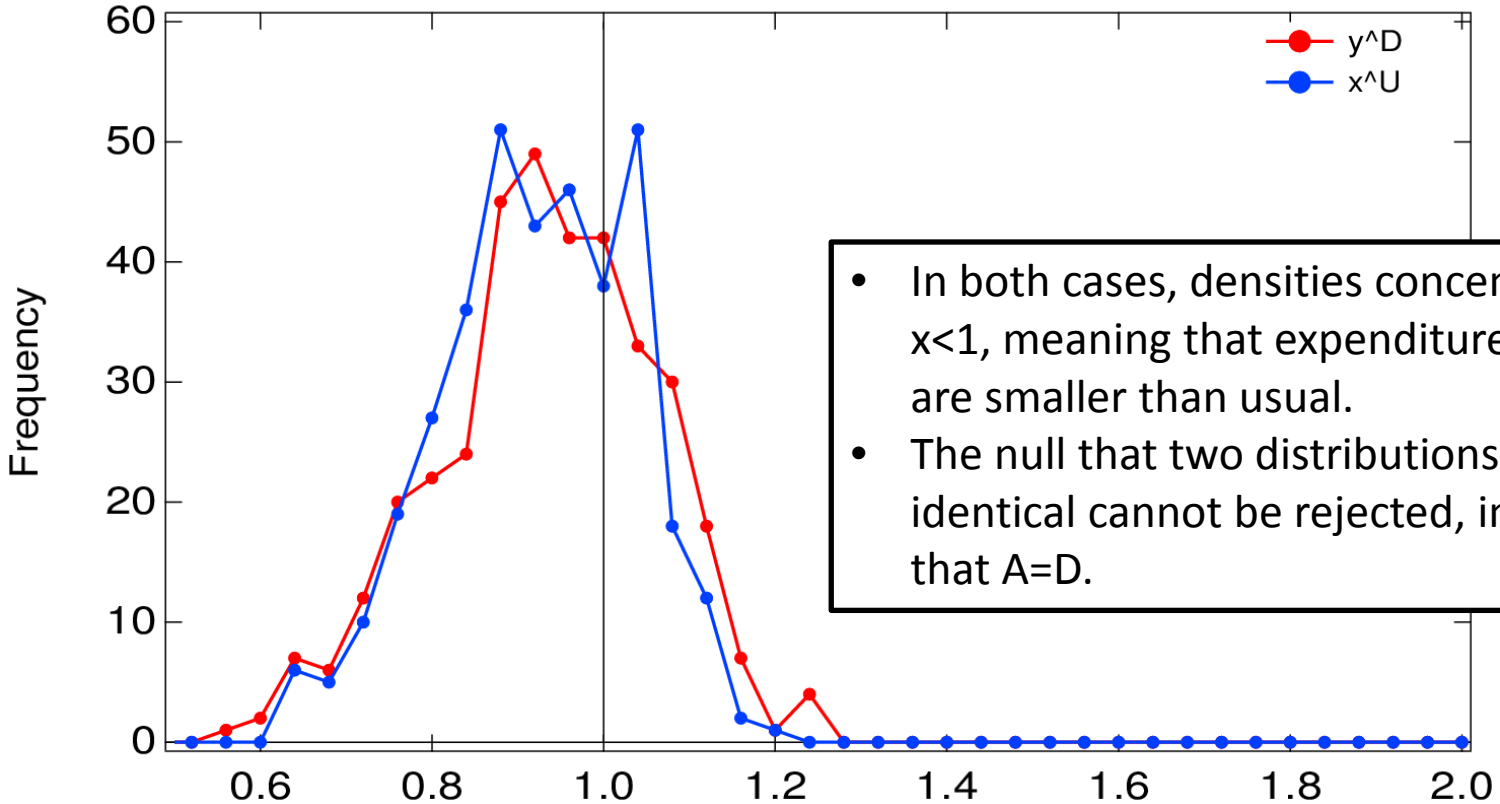
B

A

D

C

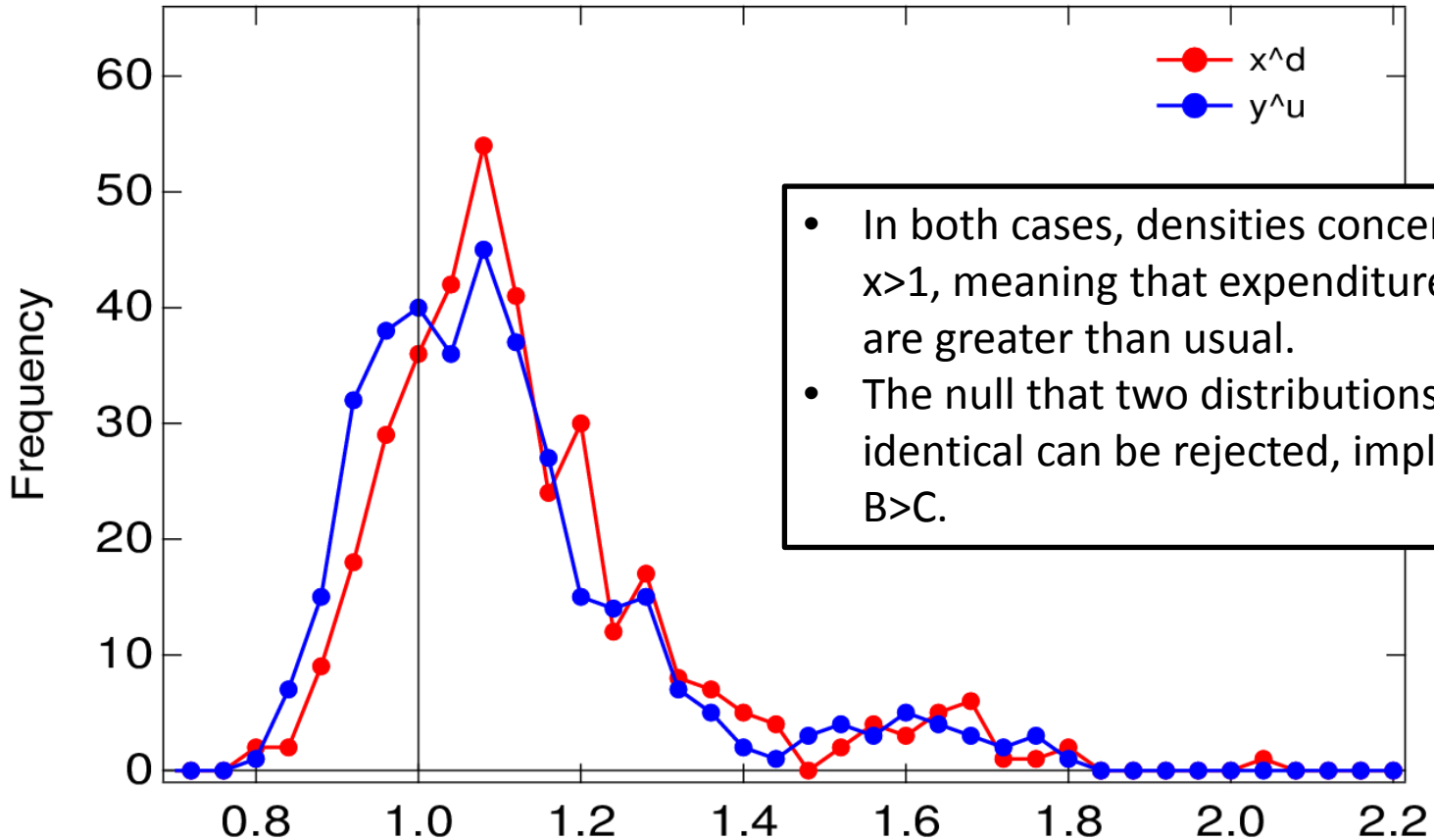
The distributions of the expenditure shares on the day before a sale period (**Red**) and on the day after a sale period (**Blue**)



- In both cases, densities concentrate at x<1, meaning that expenditure shares are smaller than usual.
- The null that two distributions are identical cannot be rejected, implying that A=D.

	Mean Welch Test	Median Wilcoxon Test	Distribution KS Test
p-value	0.290	0.248	0.170

The distributions of the expenditure shares on the first day of a sale period (**Red**) and on the final day of a sale period (**Blue**)



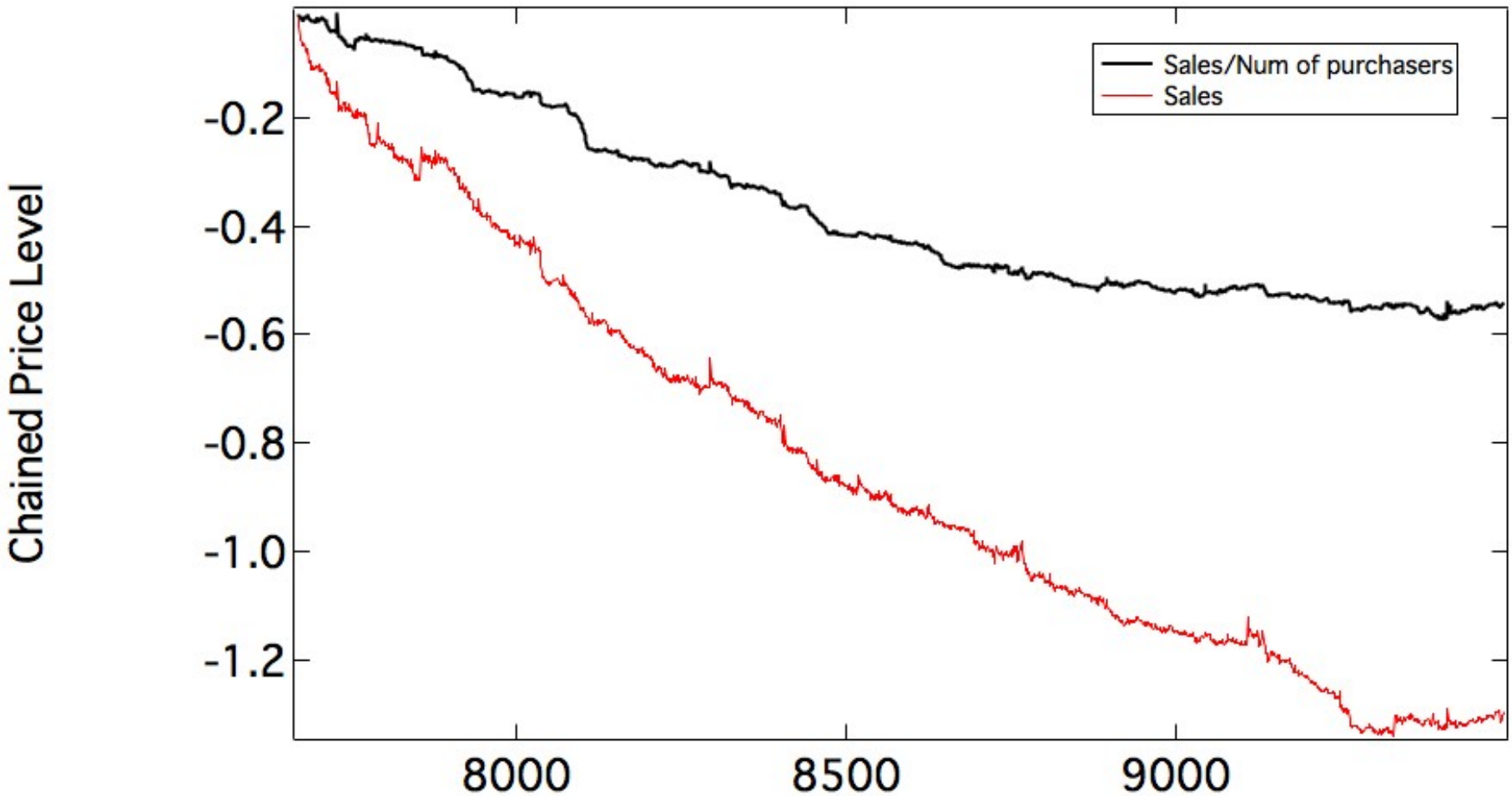
- In both cases, densities concentrate at x>1, meaning that expenditure shares are greater than usual.
- The null that two distributions are identical can be rejected, implying that B>C.

	Mean Welch Test	Median Wilcoxon Test	Distribution KS Test
p-value	0.057*	0.008***	0.016**

How to eliminate chain drift?

- Why do sales fluctuate before, during, and after a sale?
  - The price elasticity of demand for a product is determined by two margins (Levin and Yun 2011)
    Intensive margin: the quantity purchased by each individual customer increases during a sale period
    Extensive margin: the number of consumers purchasing that product at that retailer increases during a sale period

- We need to eliminate sale fluctuations due to the extensive margin. To do so, we replace sales with **sales per customer** (or sales per transaction).
  - We divide the sales of a particular product at a particular retailer on a particular day by the number of consumers who purchased the product at the retailer on that day.
  - To do so, we use the **receipt data** (i.e., the data contains information on individual purchases).


Sales/Num of purchasers
Sales
Chained Price Level
-0.2
-0.4
-0.6
-0.8
-1.0
-1.2
8000
8500
9000

Our findings on chaining

- If goods are not storable, intertemporal utility maximization implies: (1) the expenditure share of a product before a sale period is the same as that after a sale period; (2) the expenditure share on the first day of a sale period is the same as that on the final day. This symmetry implies that the Tornqvist weight associated with the price decline at the start of a sale is the same as the Tornqvist weight associated with the price increase at the end of a sale. Therefore, no chain drift arises.
- However, this no longer holds if goods are storable. In that case, the symmetry may be violated and chain drift may arise.
- Our empirical results indicate that the expenditure share of a product before a sale is the same as that after a sale.
  - This is in sharp contrast with the arguments made by previous studies including Ivancic, Diewert, and Fox (2011), and De Haan and van der Grient (2011).
- However, we find that the expenditure share on the first day is greater than that on the final day, creating a substantial downward drift.
- We find that replacing sales with sales per customer contributes to eliminating the chain drift at least partially.