"Quality Adjustment at Scale: Hedonics vs. Exact Demand Based Price Indices"

RESET Project:

Re-Engineering Statistics using Economic Transactions

Matthew D. Shapiro, University of Michigan

FCSM, November 2021

Authors: Gabriel Ehrlich(Michigan), John C. Haltiwanger (Maryland), Ron Jarmin (Census), David Johnson (Michigan), Ed Olivares (Maryland), Luke Pardue (Maryland), Matthew D. Shapiro (Michigan) and Laura Zhao (Maryland and Bank of Canada). These slides also draw from "Re-Engineering Key Economic Indicators" by Ehrlich, Haltiwanger, Jarmin, Johnson and Shapiro (forthcoming in *Big Data for 21st Century Economic Statistics*). Other researchers on RESET that have contributed to the results presented here include Seula Kim (Maryland) and Michael Navarette (Maryland).

Acknowledgements and Disclaimers

This research is supported by the Alfred P. Sloan Foundation with additional support from the Michigan Institute for Data Science, the Michigan Institute for Teaching and Research in Economics and the U.S. Census Bureau.

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

This presentation also uses data from NPD housed at the U.S. Census Bureau. All results using the NPD data have been reviewed to ensure that no confidential information has been disclosed (CBDRB-FY19-122, CBDRB-FY21-074). Opinions and conclusions expressed are those of the authors and do not represent the view of the U.S. Census Bureau.

In addition to those named on the first slide, we thank Jamie Fogel, Tian Gao, Diyue Guo, Dyanne Vaught, and Dawn Zinsser for their contribution to this research project.

We also thank staff at the U.S. Census Bureau working on related projects for their collaboration and assistance. We particularly thank Nick Orsini, Catherine Buffington, Rebecca Hutchinson and Stephanie Studds.

RESET Project Aims and Challenges

Aim:

- Re-engineer economic statistics by aggregating item-level transactions
- Consistently measure value, volume, and price
- Improve the timeliness and granularity of economic statistics

Challenge addressed in this paper:

- Rapid turnover of items requires quality adjustment at scale
- This paper examines two approaches
  - Hedonics at scale
  - Demand-based price indexes

Roadmap

- NPD results: General merchandise including online
  - Hedonics based on curated attributes/econometrics
  - Demand-based approaches
- Nielsen/Kilts: Grocery stores, pharmacies, and convenience stores
  - Machine learning hedonics based on text fields/neural network
- Not in this presentation: Work with companies data directly

Main messages

- Superlative price indices (e.g., Tornqvist, Fisher) can be readily computed in real time.
- Quality adjustment at scale is needed across a wide range of goods from high tech consumer goods to food items.

Price indices with item-level P, Q, and attribute data

- Computation of traditional superlative price indices such as Fisher and Tornqvist with t-1 and t weights at high frequency is readily feasible.
- Hedonic methods can be used to track quality change in the presence of product turnover.
- Demand-based approaches such as Feenstra (1994) adjusted Sato-Vartia and CUPI from Redding and Weinstein (2020) offer alternative methods for quality-adjusted prices.
- We compare and contrast all of these methods.

Hedonics at scale with item-level transactions data

- We consider a variety of hedonic specifications using item-level data
  - Follow several authors including: Bajari and Benkard (2005), Erickson and Pakes (2011), Byrne, Sichel and Aizcorbe (2019), and Bajari et. al. (2021)
- We focus on comparing and contrasting full-imputation hedonic method using Erickson and Pakes (2011) methodology with time dummy method

More Details of Hedonic Methods

1. Hedonic (full) imputation following Erickson and Pakes (2011) (EP) taking into account unobservable characteristics

   Step 1: Estimate levels model, log linear with time-specific coefficients:

$$\ln p_{kt} = h_t(Z_k) + \eta_{kt}$$

   Step 2: Estimate first differences with lagged residual from step 1:

$$\Delta \ln p_{kt} = Z_k' \beta_t + \kappa \hat{\eta}_{kt-1} + \nu_{kt}$$

   Use full imputation method w/Tornqvist wts to calculate "Hedonic Tornqvist, TV"

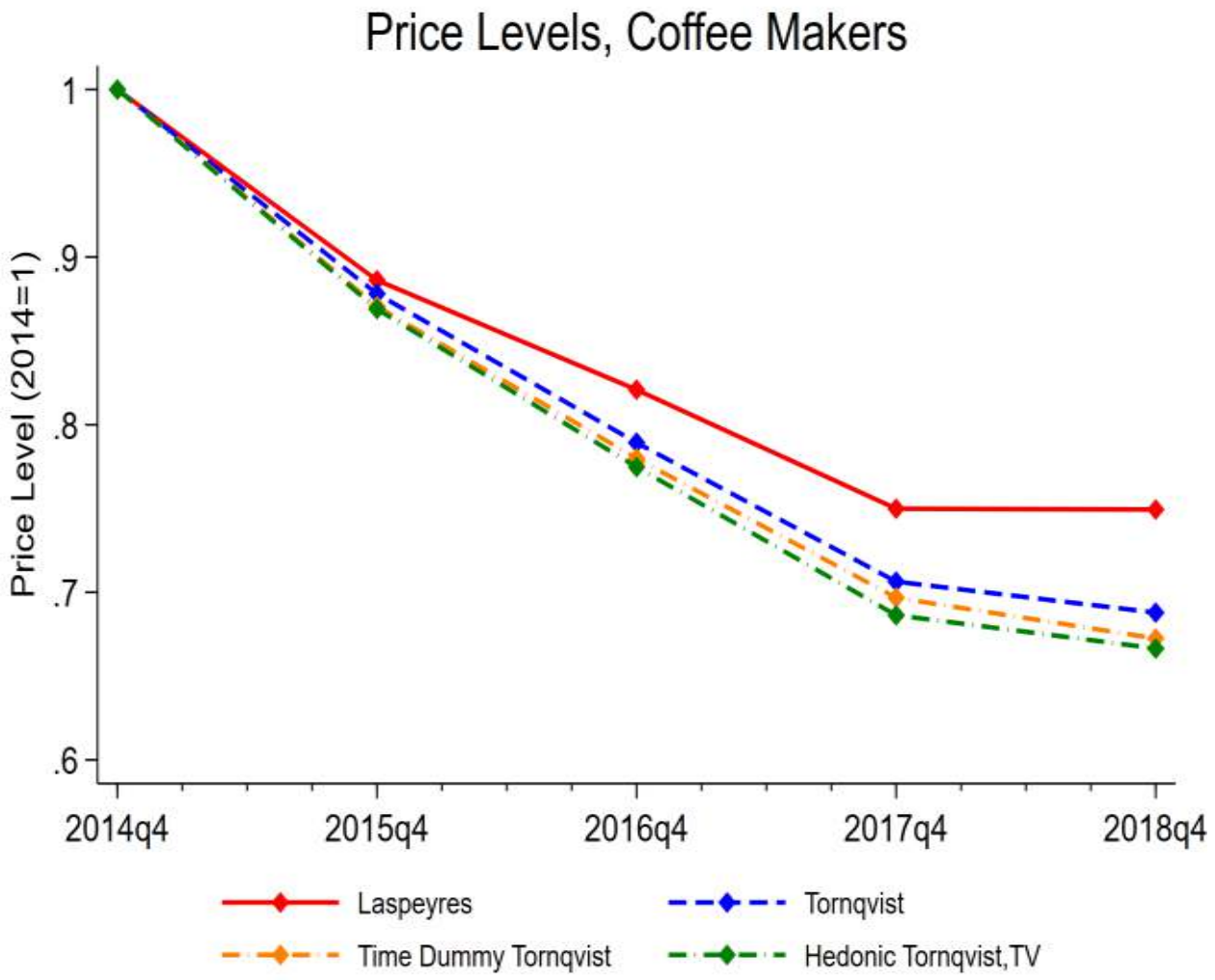
2. Hedonic Time Dummy Method
   - Use Tornqvist wts. in estimation
   - For each pair of adjacent periods, estimate:

$$\ln p_{k\tau} = \alpha_{t-1,t} + \delta_t + Z_{k\,t}' \gamma_{t-1,t} + \epsilon_{k\tau}, \qquad \tau = \{t-1, t\}$$

Demand Based Indices (CES)

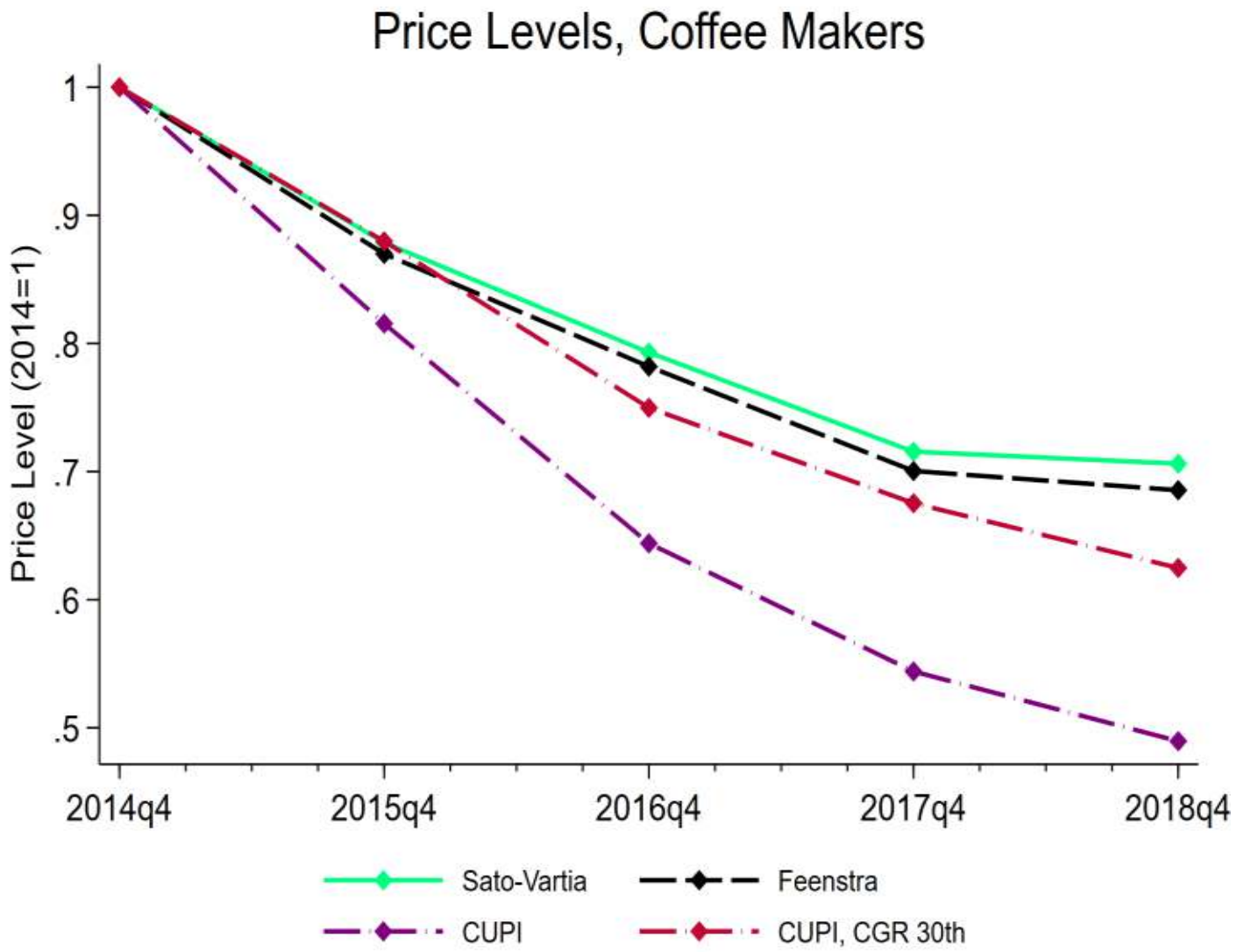
- **Sato-Vartia** is exact without product turnover and no change in relative product appeal within products over time.
- **Feenstra (1994)** adjusted Sato-Vartia is exact with product turnover and no change in relative product appeal within products over time.
- **CES Unified Price Index (CUPI) (Redding and Weinstein**, 2020) is exact with product turnover and changes in relative product appeal within products over time.
  - Common good rule important; How to treat new and exiting goods with low shares

Focus initially on one product group to illustrate findings and issues: Coffee Makers



- High pace of product entry and exit
  - 5.7% entry, 4.5% exit per qtr
- Rapid quality change
  - Single-serve pod makers entered over our sample period
- Laspeyres shows more inflation than Tornqvist (one advantage of P & Q data is ideal indices easy to compute).
- Time dummy method yields additional adjustment.
- Hedonic Tornqvist, TV lower than Time Dummy.
- All price indices are chained, quarterly.

Demand Based Indices



Rank ordering of demand based price Indices consistent with hypothesized quality adjustment:
Sato-Vartia > Feenstra > CUPI
(useful to compare within CES based indices)
However, CUPI (baseline) yields Implausibly low inflation.

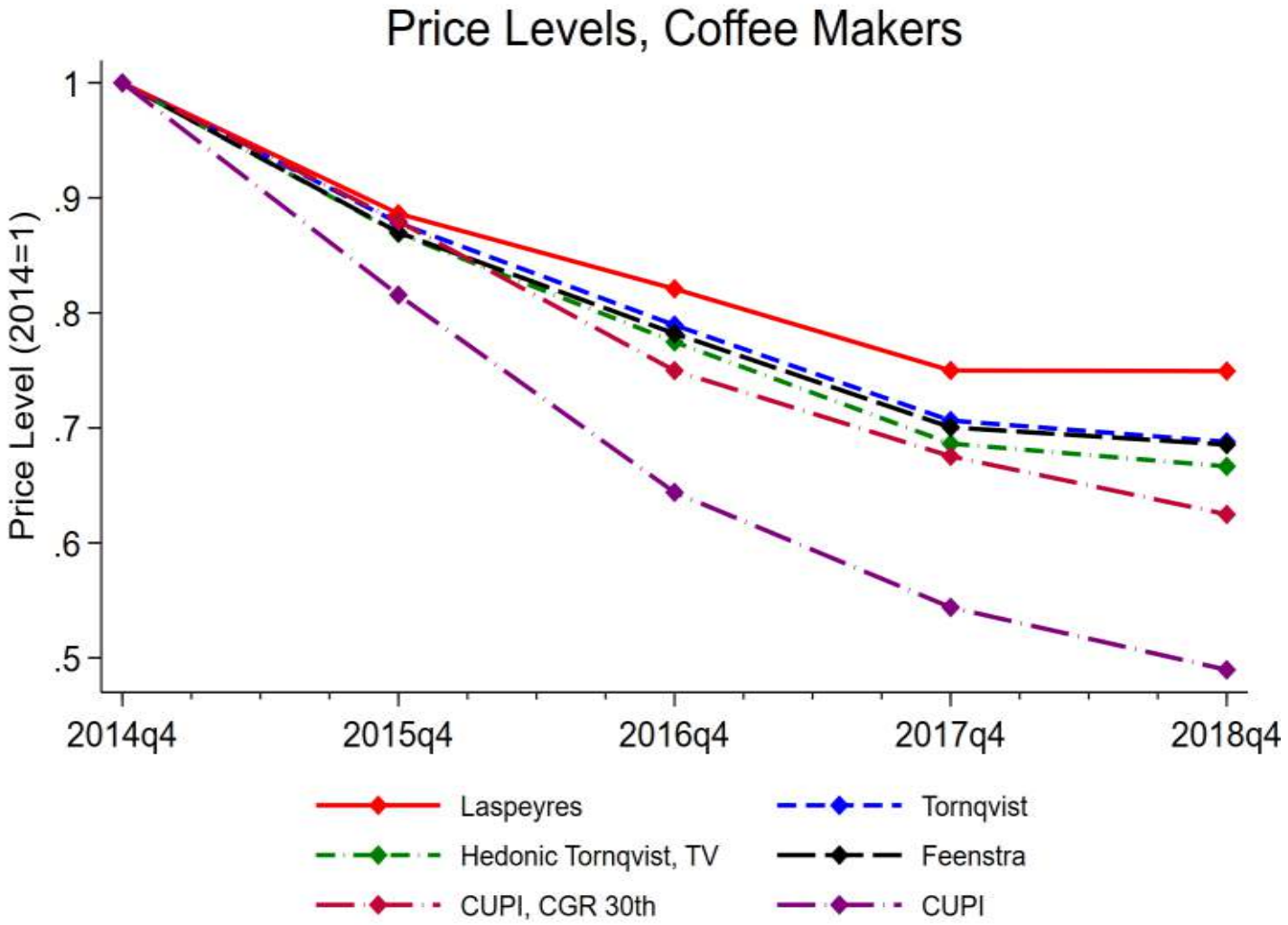
Redding and Weinstein (2020) (RW) propose an adjustment based on hypothesis that it takes times for products to enter/exit.

Here CGR reallocates goods with less than 30th percentile market share in (t-1,t) at quarterly frequency to be part of entry/exit (Feenstra) adjustment rather than Jevons or Consumer Valuation Adjustment (CVA)

Jevons and CVA are unweighted geometric means and very sensitive to small shares.

CGR rule plausible but needs more research. We include robustness analysis to alternative CGR in paper including replication of RW.

Putting pieces together – Hedonics vs. Demand Based Quality Adjusted Price Indices



All quality adjusted methods yield lower price levels than Laspeyres.

Gap with Tornqvist is smaller for Feenstra.

Both Hedonic Tornqvist, TV and CUPI imply substantial quality adjustment missing in official data.

Taking stock from NPD Results

- Need for quality adjustment pervasive
  - Paper has many goods beyond coffee makers
- Most robust method is Erickson and Pakes (2011) hedonic full imputation method incorporating time varying valuation of unobservable characteristics.
- Feenstra adjusted Sato-Vartia uniformly greater declines in price levels than Sato-Vartia (useful to compare within CES based indices)
- CUPI yields even greater declines in price levels, but sensitive to common goods rule (CGR).

Kilts/Nielsen scanner data

- Comprehensive coverage of food and non-food sold
  - Grocery stores and pharmacies
- Lacks curated product attributes in NPD data
- Can we do hedonics at scale using machine learning

Nielsen/Kilts data has marketing text descriptions, not well-coded attributes (unlike NPD)
Useful Challenge!

Soft drinks examples:

‘brand’ ZR DT LN/LM CF NBP CT

‘brand’ NATURAL R CL NB 12P

Toilet paper examples:

‘brand’ DR W 1P 308S TT 6PK

(specific brands suppressed)

Could hire a team to code:

Soft drinks examples:

‘brand’ ZR DT LN/LM CF NBP CT

‘brand’ NATURAL R CL NB 12P

DT=diet, R=regular, 12P=12 pack

Toilet paper examples:

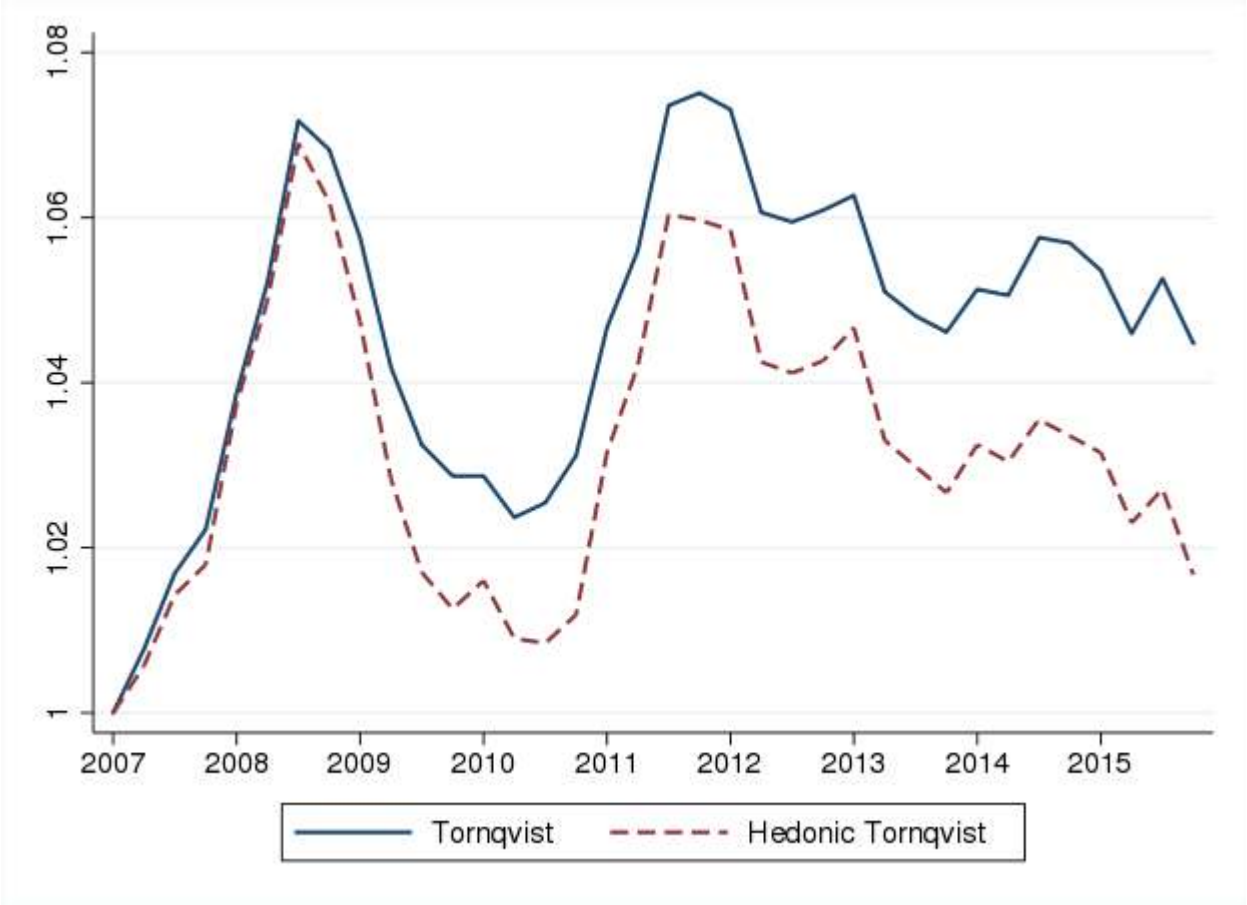
‘brand’ DR W 1P 308S TT 6PK

1P= 1 ply, 308S= 308 sheets, 6PK=6 pack

Use Machine Learning (ML) – integrated with Erickson and Pakes (2011)

- We work with team of computer scientists at Michigan and MIT led by Mike Cafarella.
- Use neural network to create mapping between prices and characteristics for Nielsen using abbreviated text fields and brands.
- Bajari et. al. (2021) use neural net model for hedonics
  - Innovation in Bajari et al (2021) is ML with text fields and images.
- Our innovation: Integrate Erickson and Pakes into ML
  - First create predicted prices in levels. Generate residual
  - Repeat ML with lagged residual as additional embedding

Food, Nielsen/Kilts Data

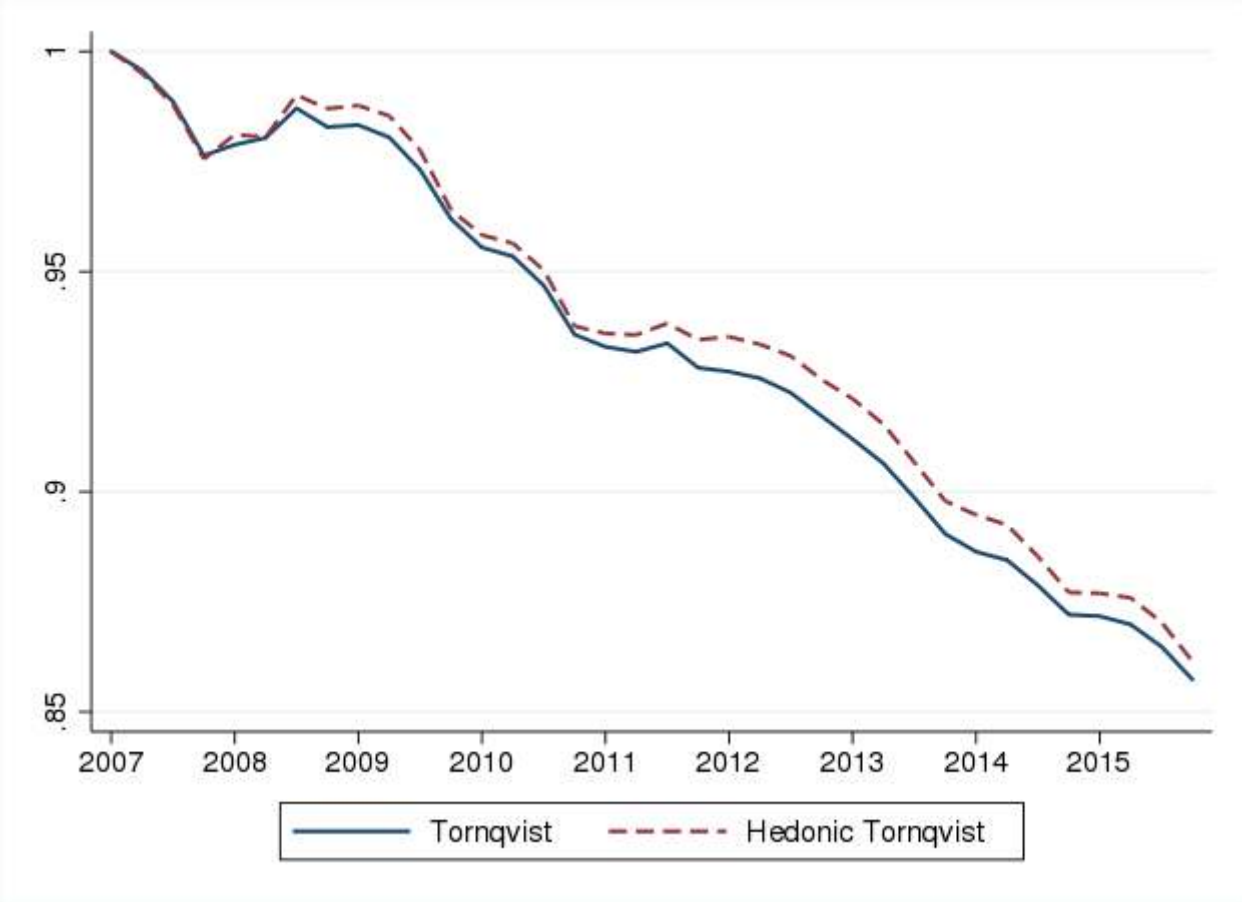


BLS CPI and Nielsen Laspeyeres track each other closely (not reported here)→Nielsen and BLS have similar price

Substantial quality change—even in food

- Product turnover substantial
- Quality change in food substantial
- ML techniques effective!

Non-Food, Nielsen/Kilts Data



No quality change found in non-food

***Not* evidence of lack of quality change**

1. Exit of non-food goods from grocery stores, not from the economy
2. ML hedonics show, *unusually*, exiting goods are higher quality

Deconstructing Effects of Entry and Exit on Quality-Adjusted Price Index

Exit adjustment

Entry adjustment

Net adjustment

Table 1: Cumulative Inflation Rate, 2007Q1 to 2015Q4

	Food		Nonfood	
	Cumulative	Quarterly	Cumulative	Quarterly
Laspeyres	13.4%	0.33%	-15.3%	-0.44%
Hedonic Laspeyres	11.1%	0.28%	-11.7%	-0.33%
Paasche	-3.7%	-0.10%	-13.2%	-0.37%
Hedonic Paasche	-6.9%	-0.19%	-16.0%	-0.46%
Tornqvist	4.5%	0.12%	-14.3%	-0.40%
Hedonic Tornqvist	1.7%	0.04%	-13.9%	-0.39%

Deconstructing Effects of Entry and Exit on Quality-Adjusted Price Index

For Non-Food:
Exiting goods are of higher quality than continuing good

Table 1: Cumulative Inflation Rate, 2007Q1 to 2015Q4

	Food		Nonfood	
	Cumulative	Quarterly	Cumulative	Quarterly
Laspeyres	13.4%	0.33%	-15.3%	-0.44%
Hedonic Laspeyres	11.1%	0.28%	-11.7%	-0.33%
Paasche	-3.7%	-0.10%	-13.2%	-0.37%
Hedonic Paasche	-6.9%	-0.19%	-16.0%	-0.46%
Tornqvist	4.5%	0.12%	-14.3%	-0.40%
Hedonic Tornqvist	1.7%	0.04%	-13.9%	-0.39%

Exit adjustment

Entry adjustment

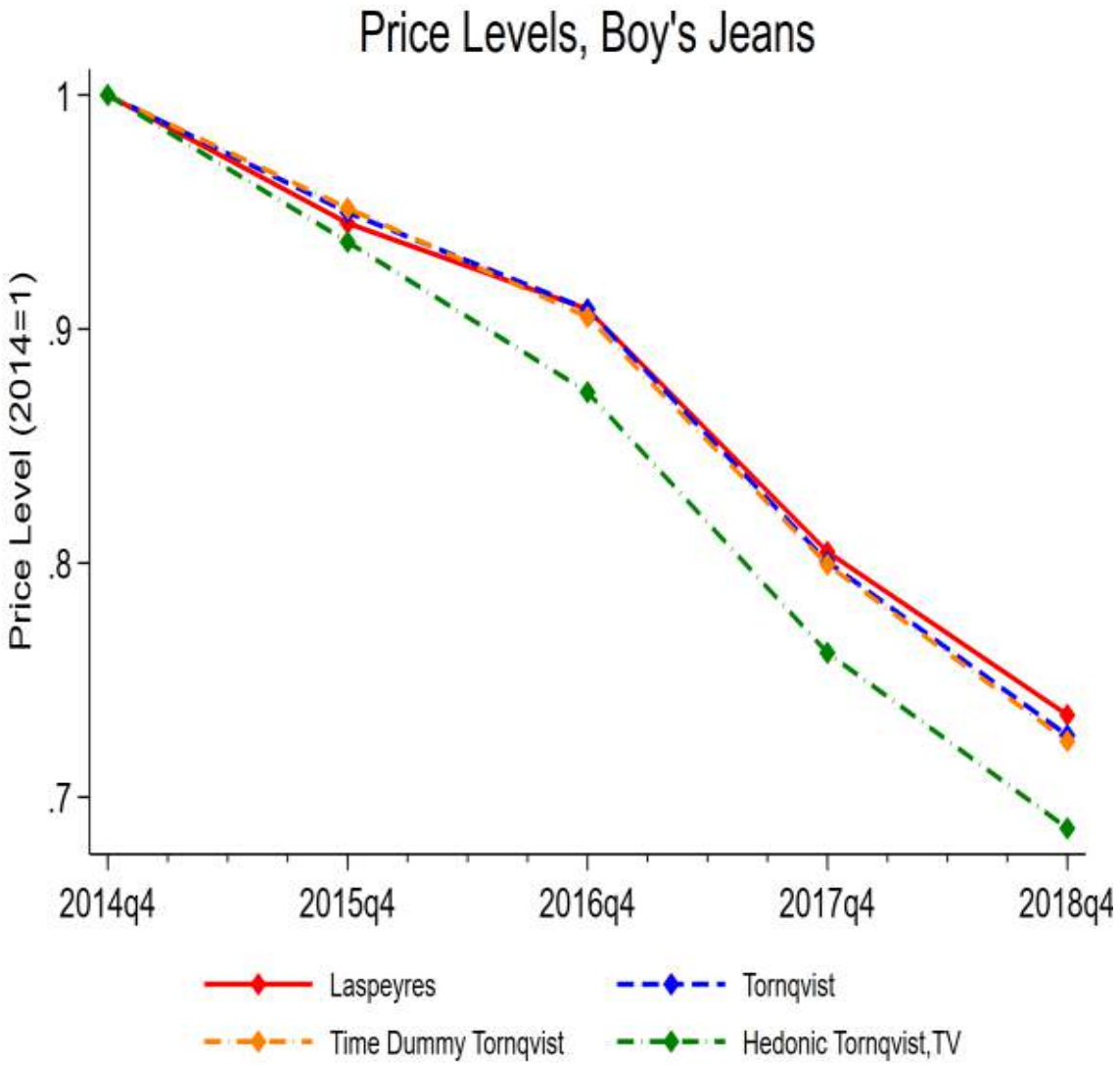
Net adjustment

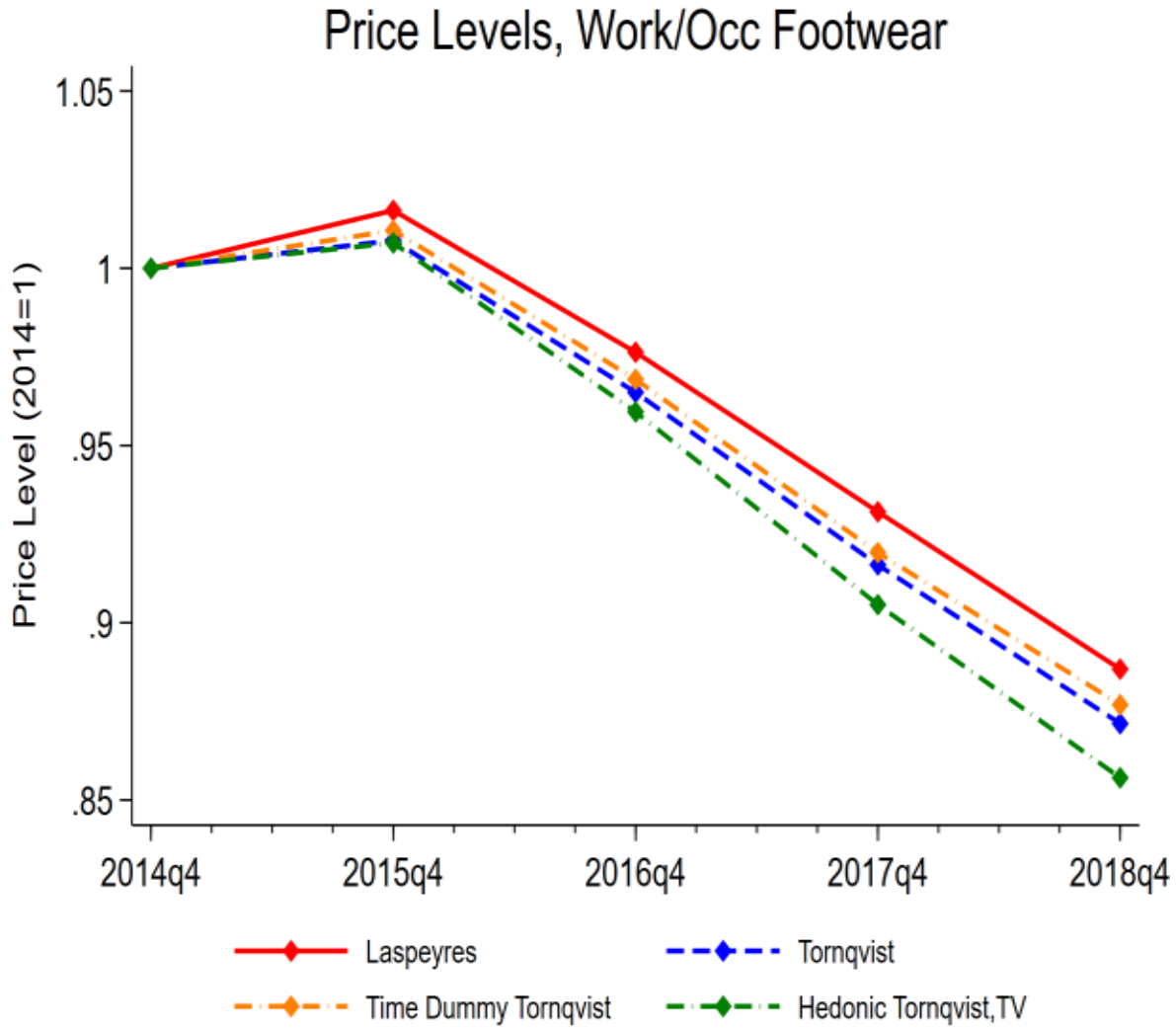
Taking stock

- Item-level P and Q data can be used to produce internally consistent nominal sales and price deflators that adjust for quality.
- Most robust results for quality-adjusted prices are using hedonics with econometric methods that account for time varying unobservables as well as observables.
  - Machine learning on product descriptions in terms of text and images facilitates using these methods at scale and overcoming disparate product attribute data.
  - Integration of machine learning approach with econometrics insights is an important innovation.
- Demand theory approach using CES approach is readily implementable at scale with these data. Should be in the toolkit. However, key limitations include
  - (i) defining common goods; (ii) defining CES nests; (iii) estimating elasticities; and (iv) more generally specification error (strong assumptions underlie implementation).

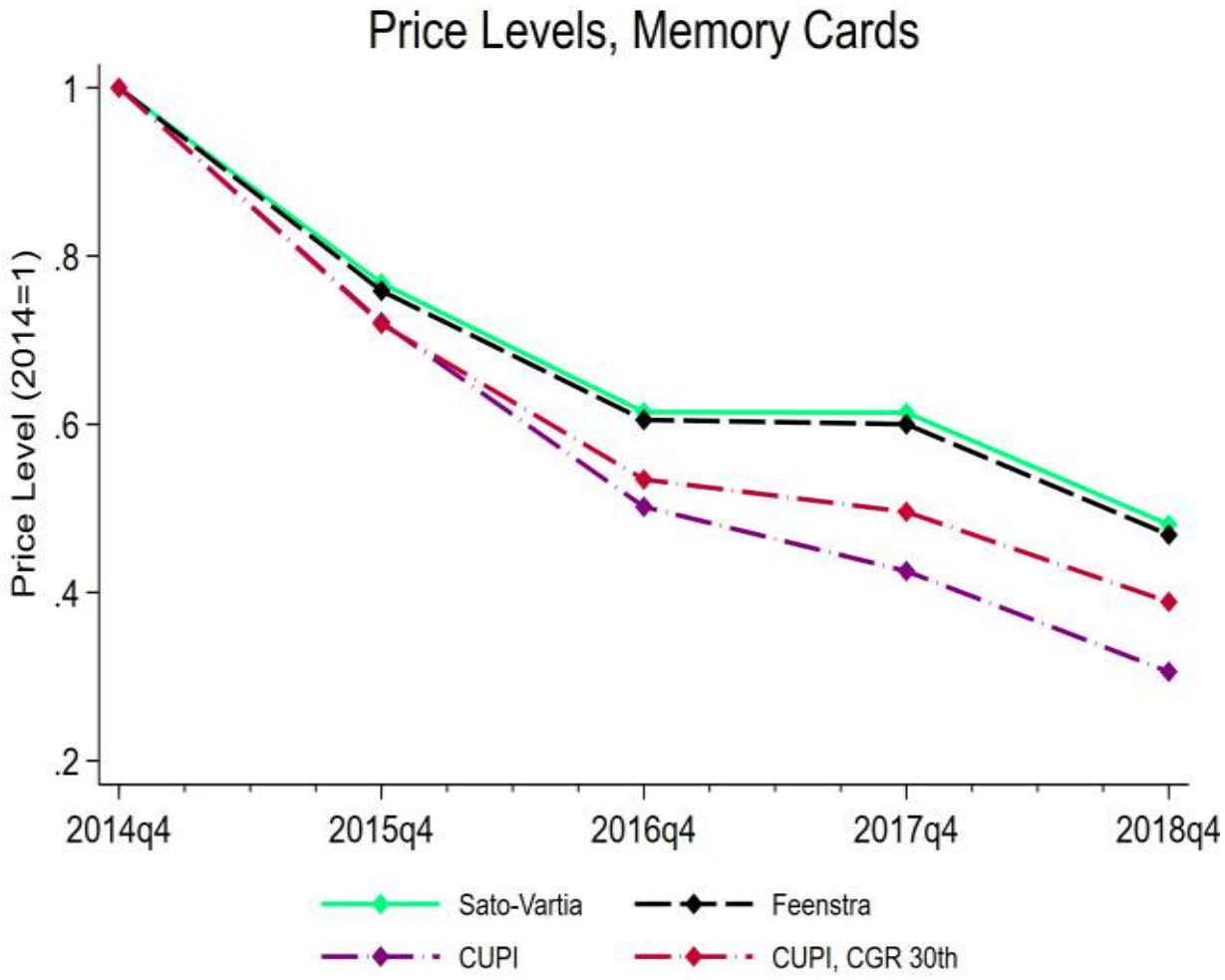
Additional Slides

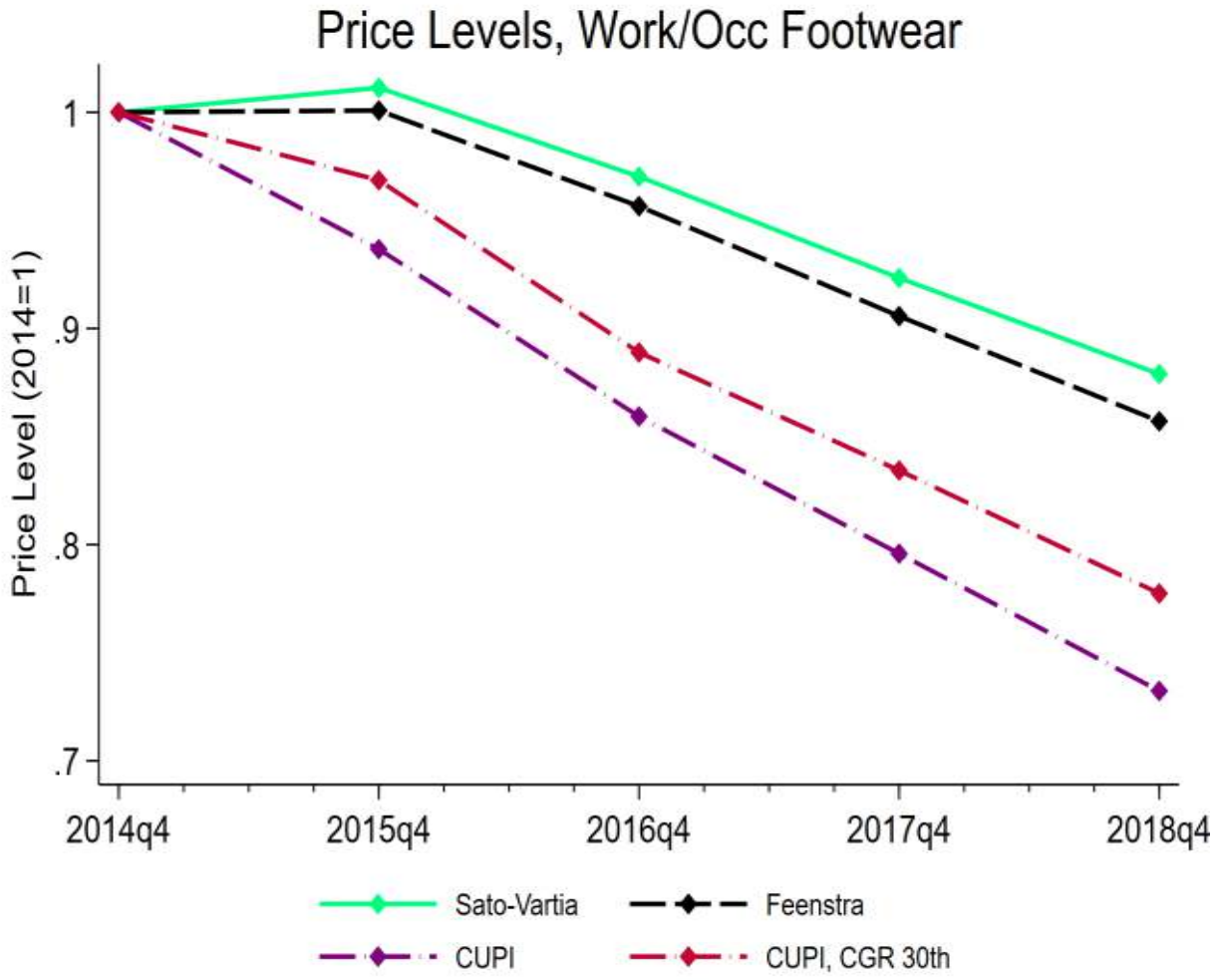
Hedonics: For boys jeans and footwear



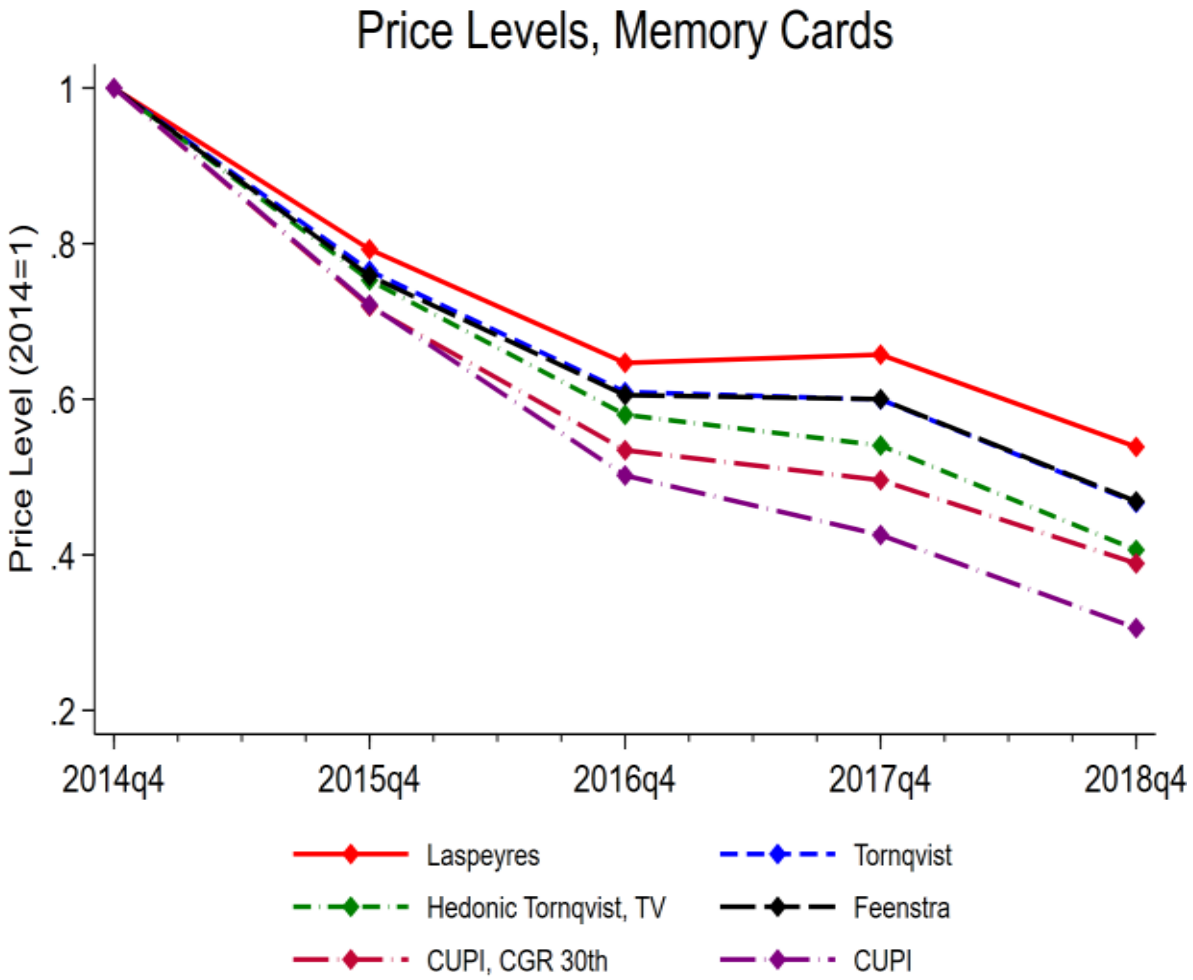


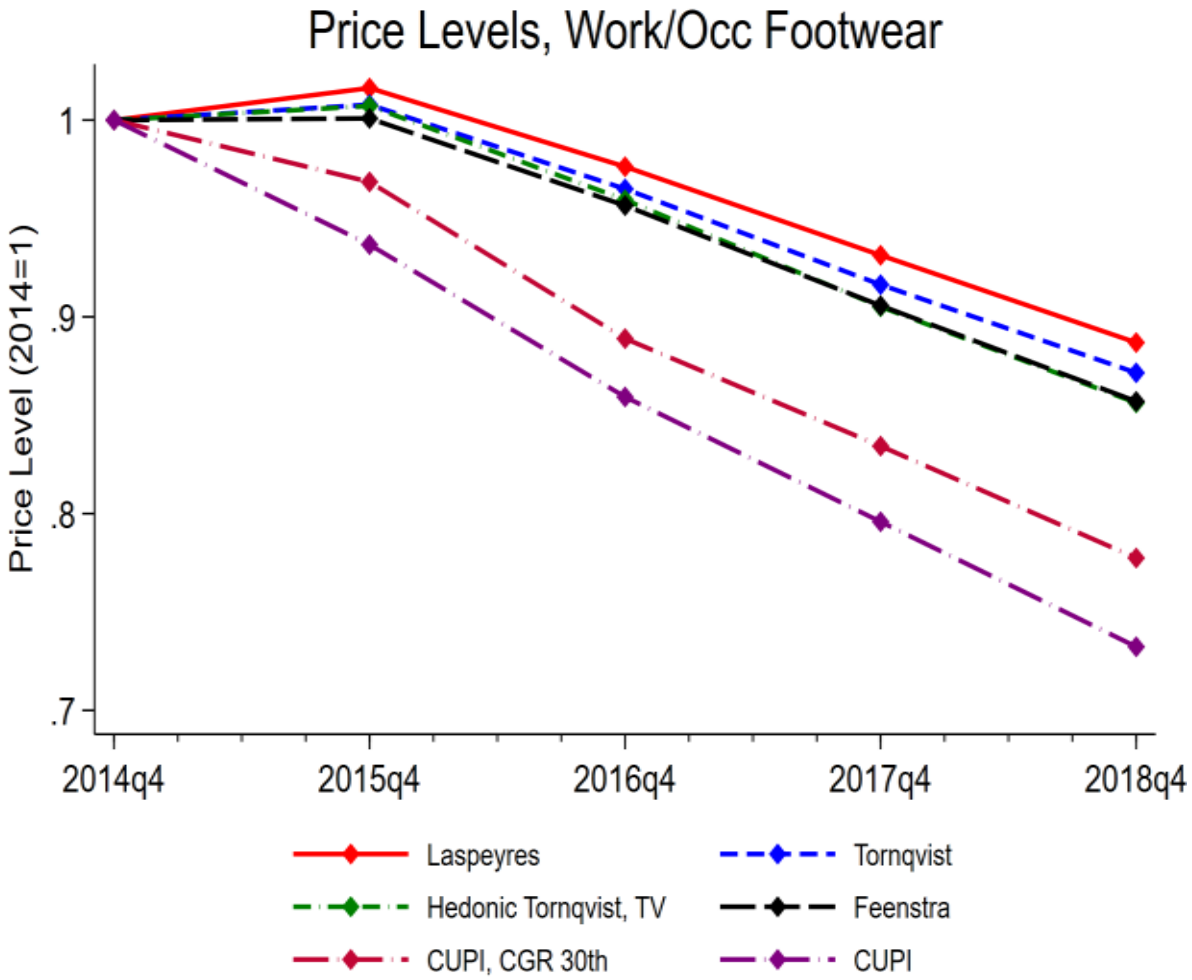
Demand Based Indices: Memory Cards and Work/Occ Footwear





Hedonics vs Demand Indices: Memory Cards and Work/Occ Footwear





Cuts for length from FCSM

Statistical Agencies Moving into 21st Century

- Our system of official economic statistics was mostly developed in the mid-20th Century
- Separate surveys for key concepts across different agencies
  - Census: Revenue and Expenditures
  - BLS: Prices, Wages, and Employment
  - BEA: Integration of data to build national accounts

Surveys use different business frames across agencies

- Surveys are mostly online or mailed, but BLS uses store-level visits for CPI
- Reliance on surveys means low frequencies and limited adjustment for quality change
  - CPI and PPI weights updated every 2 and 5 years, respectively
  - Econ Census every 5 years
  - About 7 percent of goods in CPI have hedonic quality adjusted prices

Statistical Agencies Moving into 21st Century

- End-state objective: Re-engineer key economic indicators such as real output and inflation to release **consistent, timely, and granular** statistics
- Census, BLS, and BEA exploring integrated data collection from naturally occurring data
- Quality-adjusted measures of real output and inflation for all goods
- Reduced survey burden on firms
- Data harvested from item-level transactions data that firms and information aggregators are actively using already
- **The RESET project:**
  - **Address conceptual, practical, and contractual issues for implementation at scale**
  - **Blueprints for new architecture for collecting data and creating official statistics.**

Why Is Re-engineering Necessary?

- Measuring inflation central to economic policy and decision-making
  - Estimates of important unmeasured quality change
    - Large relative to movements in inflation that trigger changes in monetary policy
  - Federal budget heavily indexed to CPI inflation
    - Social Security cost-of-living adjustments
    - Tax brackets
  - Overstating inflation understates growth and productivity
    - Do we need to rethink recent slow economic performance?
- Policymakers and data users need timely and granular estimates.
- Official statistics interpolate GDP
  - Estimates are too smooth
  - Missed collapse of GDP in late 2008 owing to financial crisis
  - Change in composition of retail spending during and after the pandemic

Data sources for current project

- NPD (general merchandise stores, including online)
  - High-quality attributes at item level from value-added by NPD
- Nielsen (Kilts Center: grocery, discount, convenience, drug and liquor store items for food and nonfood)
  - Limited-quality item descriptions
- Individual retailers (working behind their firewalls)
  - Item-level prices and quantities
  - No results in today's presentation, but highlights projects broader scope

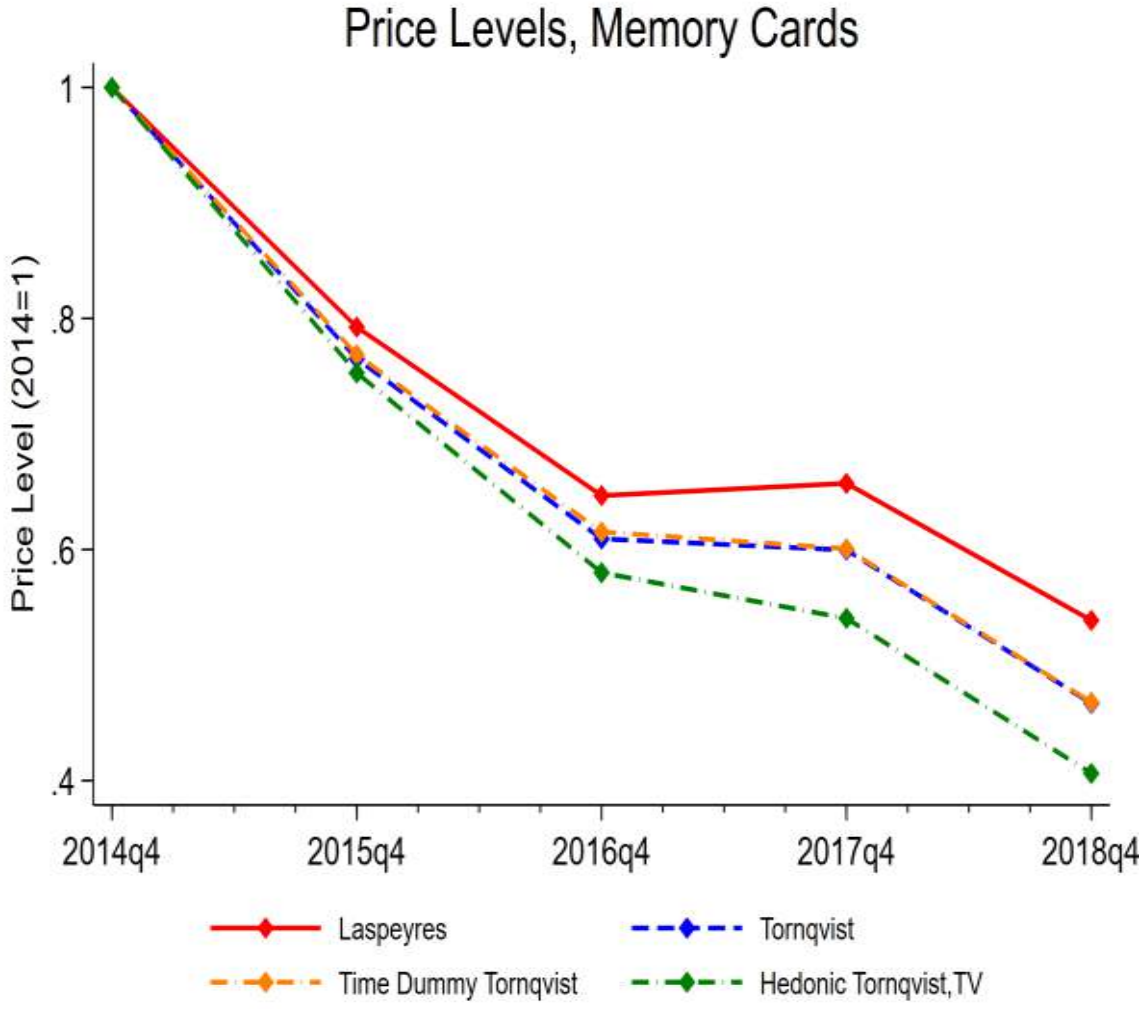
Key Challenge/Opportunity for Adjusting Prices for Quality: Enormous Product Turnover

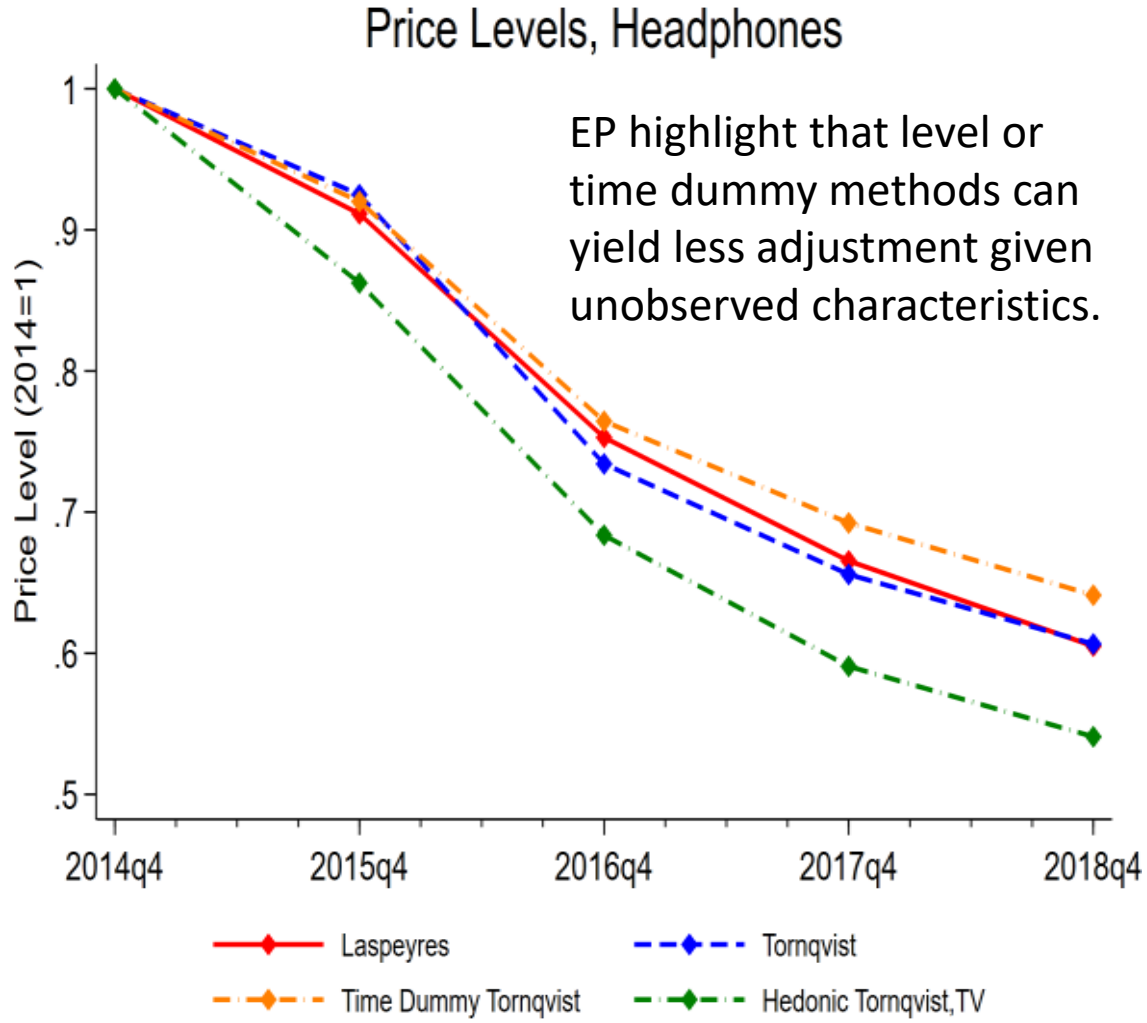
- High product (item-level) entry and exit rates
- Some turnover is substantive, some is marketing/packaging
- Traditional price indices ignore this turnover
- Hedonics using attributes data or demand based indices that take into account product turnover can account for quality improvement

CES Demand Based Indices: Formulas

- Sato-Vartia: $\ln \Phi_{t-1,t}^{SV} = \sum_{k \in C_{t-1,t}} \omega_{kt} \ln\left(\frac{p_{kt}}{p_{kt-1}}\right), \quad \omega_{kt} = \frac{s_{kt} - s_{kt-1}}{\ln s_{kt} - \ln s_{kt-1}} \Big/ \sum_{k \in C_{t-1,t}} \frac{s_{kt} - s_{kt-1}}{\ln s_{kt} - \ln s_{kt-1}}$
  - Limitation: does not allow for product turnover or appeal shocks
- Feenstra: $\ln \Phi_{t-1,t}^{SV} + \frac{1}{\sigma - 1} \ln \frac{\lambda_{t,t-1}}{\lambda_{t-1,t}}, \qquad \lambda_{t,t-1} \equiv \frac{\sum_{k \in C_{t-1,t}} P_{kt} C_{kt}}{\sum_{k \in \Omega_t} P_{kt} C_{kt}} \quad \lambda_{t-1,t} \equiv \frac{\sum_{k \in C_{t-1,t}} P_{kt-1} C_{kt-1}}{\sum_{k \in \Omega_{t-1}} P_{kt-1} C_{kt-1}}$
  - $\sigma = elasticity\ of substitution$
  - Adjusts for changing quality via product turnover
- CUPI: $\ln \Phi_{t-1,t}^{CUPI} = \frac{1}{\sigma - 1} \ln \frac{\lambda_{t,t-1}}{\lambda_{t-1,t}} + \frac{1}{N_{t-1,t}^{C}} \sum_{k \in C_{t-1,t}} \left[\ln \frac{p_{kt}}{p_{kt-1}} + \frac{1}{\sigma - 1} \ln \frac{s_{kt}}{s_{kt-1}} \right]$
  - Three terms: Feenstra adjustment, Jevons and consumer valuation adjustment (CVA).
  - CVA captures relative change in product appeal within narrow product groups.
  - $C_{t-1,t}$ is set of continuing products, Ω_t is set of all products sold at time t

For Hedonics: Hedonics Tornqvist, TV patterns robust, not Hedonic Time Dummy

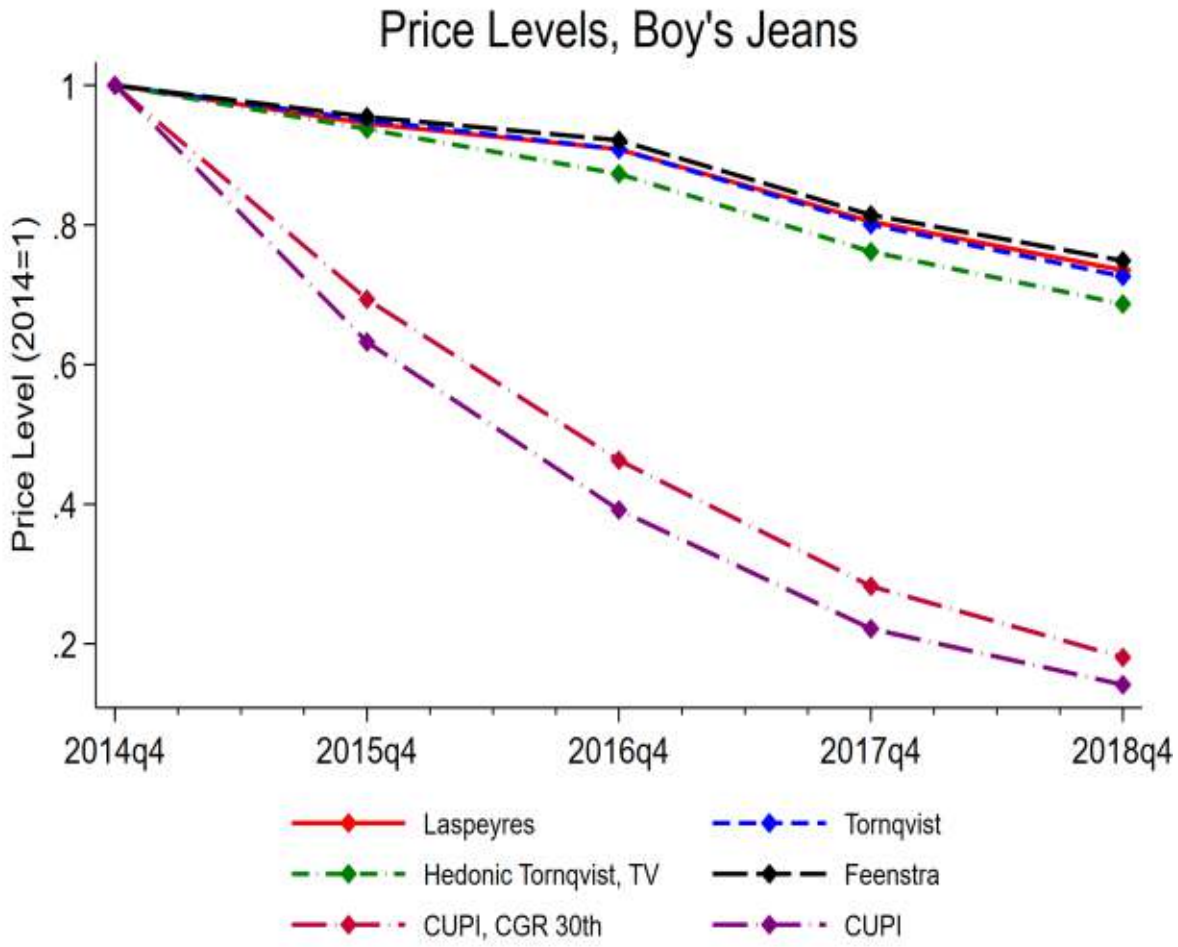


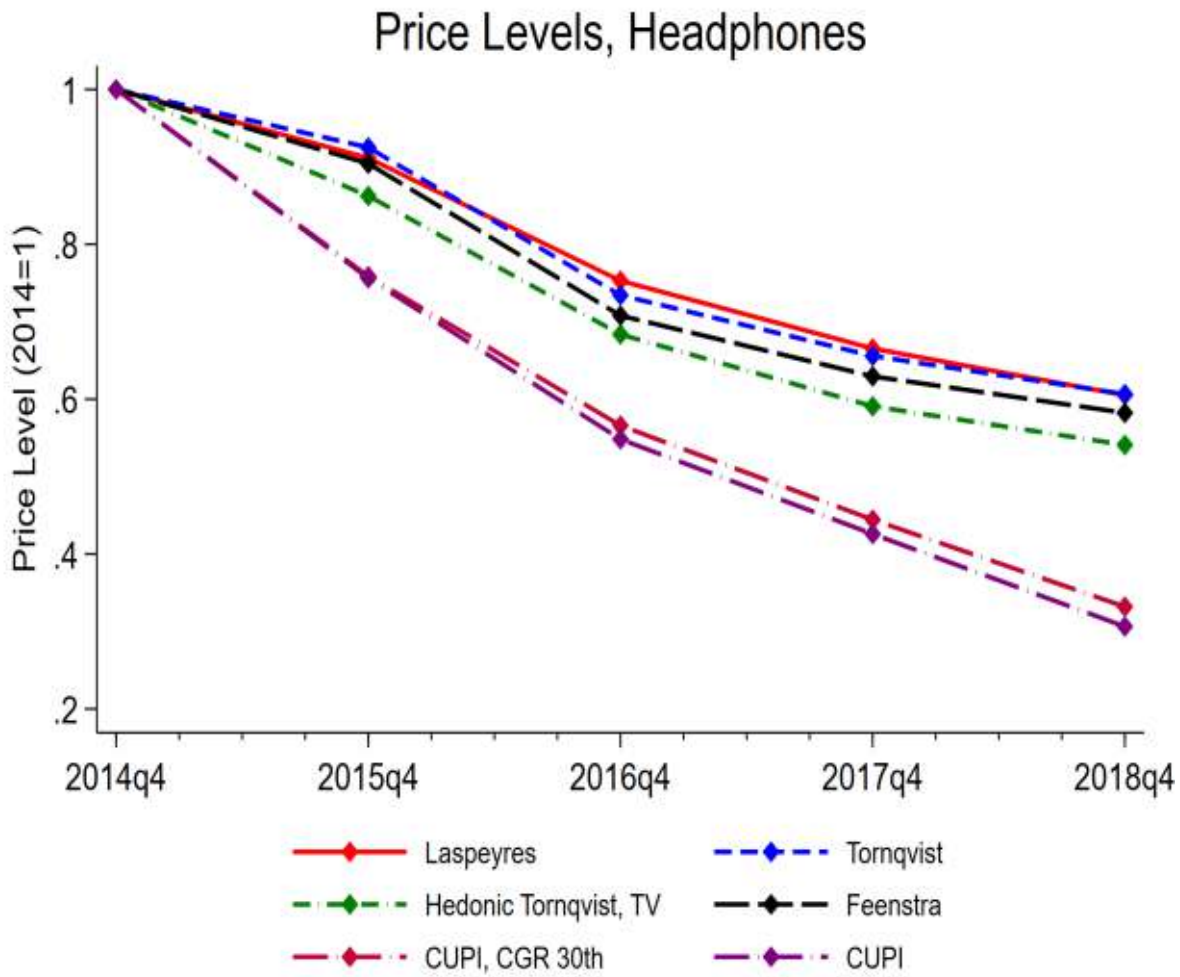


EP highlight that level or time dummy methods can yield less adjustment given unobserved characteristics.

Observe difference in scales.

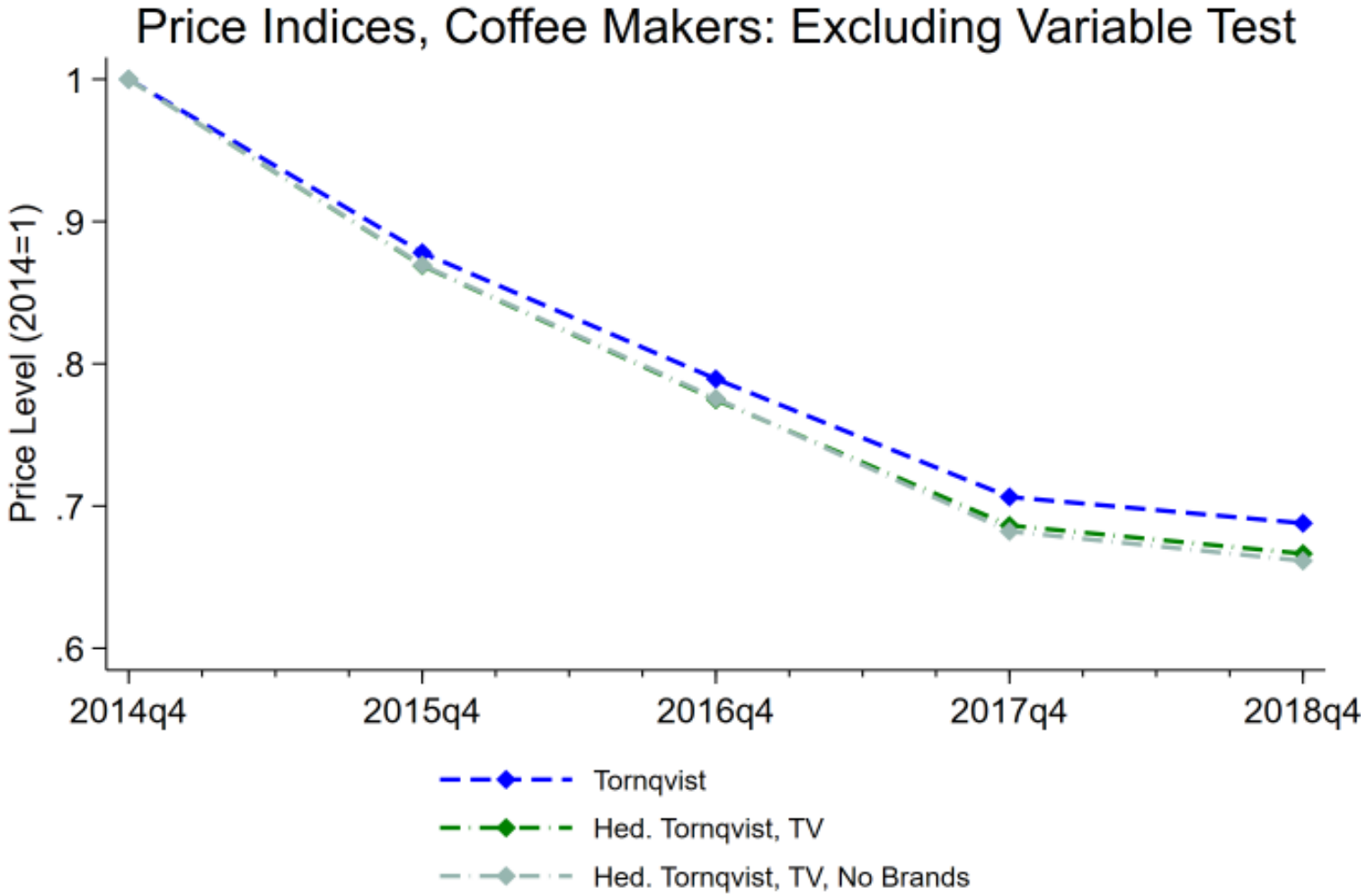
For CUPI, patterns not robust to CGR – for Headphones and Boys Jeans, even with CGR at 30th percentile, CUPI is implausibly low





Additional challenges for demand based: 1. Defining CES product group; 2. Estimation of Elasticities.

Robustness of Hedonic Tornqvist, TV (test intentionally leave out characteristics)



These findings mimic those in Erickson and Pakes (2011).

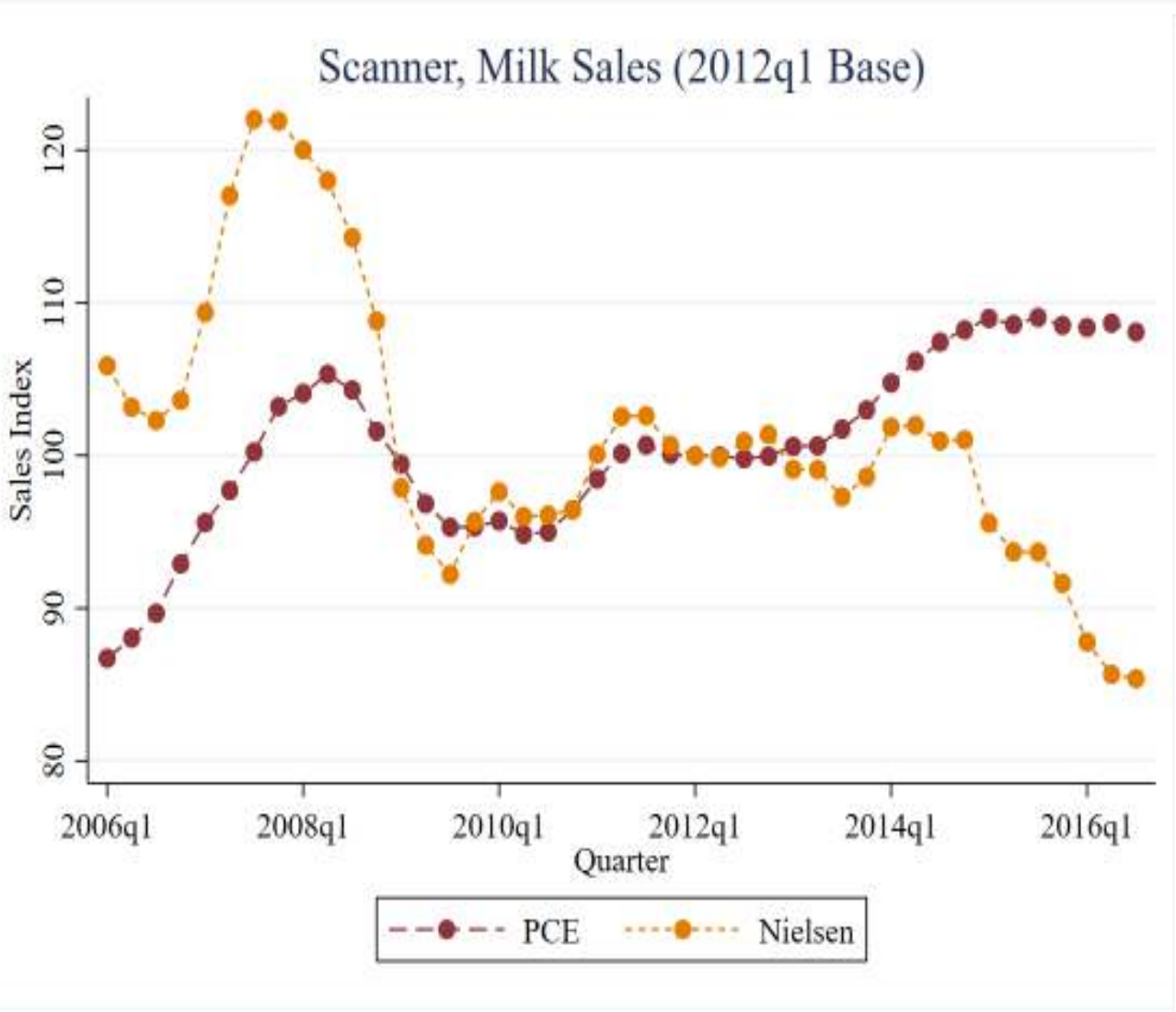
They find that their method works well with intentionally omitted characteristics for televisions.

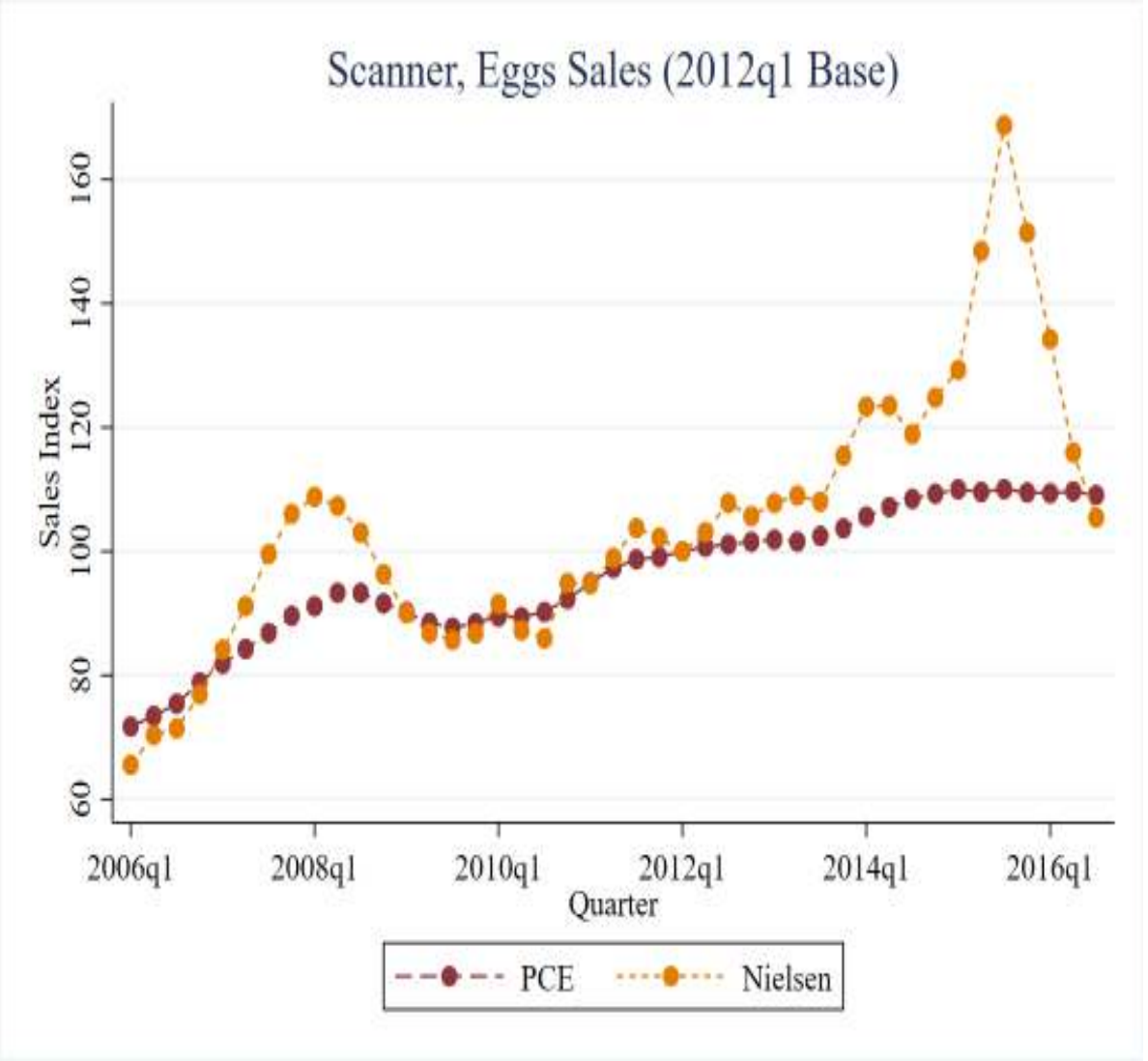
We find, like they do, that level methods are not robust to intentionally leaving out characteristics.

Final point:
Internally consistent P and Q data also important for measuring sales=P*Q

- High frequency (monthly, annual) sales from surveys (MRTS, ARTS) at the firm (not establishment) level without product detail and limited industry information.
- Product detail available every five years from Economic Census
- BEA builds PCE using MRTS, ARTS and interpolates and extrapolates from Economic Census data.
  - Supplements from other sources to improve imputation.
- Core component of this project is to compared nominal sales from PCE, Nielsen and NPD for harmonized product groups.
- Differences in nominal sales patterns + Differences in price indices both contribute to potential differences in real output patterns.

Examples of two harmonized product groups from PCE and Nielsen





Nielsen more cyclical consistent with hypothesis that PCE is potentially too smooth. Exploring post 2014: Decline in sales for traditional milk, increase in plant-based (e.g. soy) milk.

Nominal sales increased for eggs in 2015 given bird flu (spike in prices). PCE smoothing method missed that variation.