Re-Engineering Key National Economic Indicators

Gabriel Ehrlich (Michigan), John C. Haltiwanger (Maryland), Ron Jarmin (Census), David Johnson and Matthew D. Shapiro (Michigan)

Presentation at FESAC

June 2019

Acknowledgements and Disclaimers

This research is supported by the Alfred P. Sloan Foundation with additional support from the Michigan Institute for Data Science. 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). Opinions and conclusions expressed are those of the authors and do not necessarily represent the view of the U.S. Census Bureau. We thank Jamie Fogel, Diyue Guo, Edward Olivares, Luke Pardue, Dyanne Vaught, and Laura Zhao for superb research assistance.

Status quo: Decentralized data collections

Real output

- Census collects the "numerator": Revenue
- BLS collects the "denominator": Prices
- BEA does the division: $Q = P*Q/P$

Non-simultaneous collection of price and quantity

- Stratified surveys from small and deteriorating samples
- Mismatch of price and revenue data
- High cost and burden
- Difficulty of accounting for changes in products

Measuring Real and Nominal Consumer Spending—Current Architecture

Census (nominal spending)	**BLS (prices)**
Data collection: **Retail Trade surveys (monthly and annual)** **Economic Census (quinquennial)** **Consumer expenditure survey (conducted for BLS)**	**Data collection:** **Consumer Expenditure survey (used for spending weights), collected under contract by Census** **Telephone Point of Purchase survey (purchase location)[a]** **CPI price enumeration (Probability sampling of goods within outlets)**
Published statistics: **Retail Trade (monthly and annual) by firm type** **Retail Trade (quinquennial) by product class**	**Published statistics:** **Consumer Price Index (monthly) by product class**

BEA (aggregation and deflation)

Data collection:
Census and BLS data; supplemented by multiple other sources

Published statistics:
Personal Consumption Expenditure: Nominal, real, and price (monthly)
GDP (quarterly)

Reengineered data for retail P and Q

Item-level transactions data

- Item-level data allows inferring price from sales and quantities
- Price, quantity and revenue measured
  - Simultaneously
  - At high frequency
  - Universe (or large sample) of transactions
  - With little lag
  - With reduced need for revisions
  - With granular information on location of sale (geography, store/online)
  - Immediate accounting for changes in goods

Devil in the Details

Transactions data much more readily available for retail than other sectors

- Personal Consumption Expenditures (PCE) 68% of GDP
- Goods 31% of PCE
- Goods (less vehicles, fuel, and prescription drugs) 22% of PCE

Many conceptual and measurement issues need to be resolved before practical implementation in the statistical agencies

Continuity is official statistics is important

Changes in Price Index (e.g., CPI) methodology have powerful implications for policy:

- Monetary Policy: Inflation a target for policy
- Fiscal Policy: Indexation of Social Security benefits and tax brackets

Current Agency Activities: Today’s presentations

Census

- Evaluating use of point-of-sale data for measuring retail sales
- Addressing survey non-response

BLS

- Multiple sources of big data being considered for CPI
- Some sources being implemented

Not (yet) re-engineering

- Using big data to replace/supplement existing surveys/enumerations
- Not yet integrating price and quantity measurement

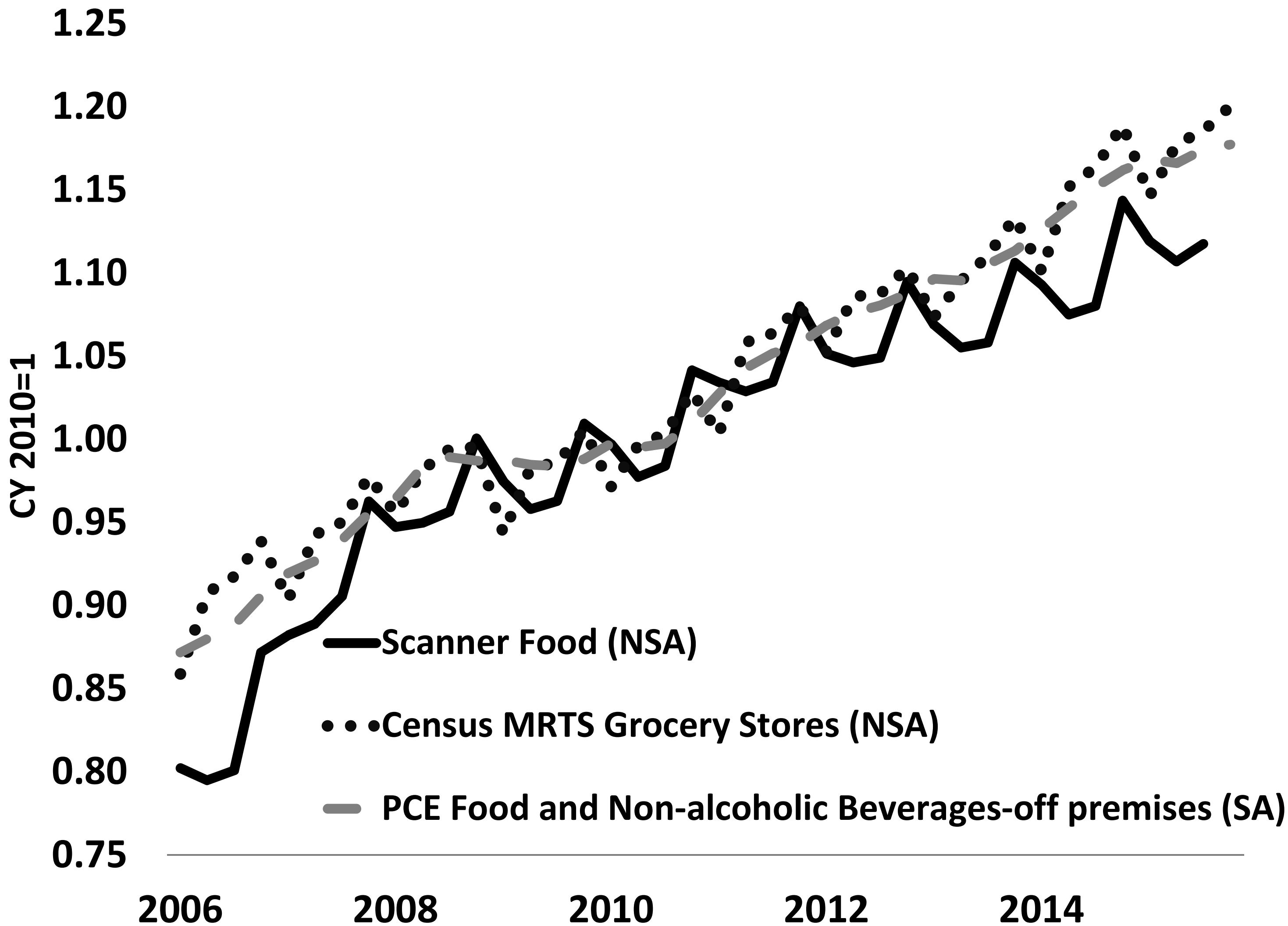
Roadmap of analysis presented today

Objective is to explore alternative methods for measuring revenue, real revenue and prices that are derived from the same (item-level transactions) source data.

Exploratory exercises using scanner data for P and Q

- Nielsen covers grocery stores and mass merchandisers
  - More than 100 product groups and 1000 product modules (millions of products).
  - Classify into Food and NonFood items
    - Food nominal expenditures: Compare scanner data to Census surveys and Personal consumption expenditures for food (Scanner provides high frequency product detail)
    - Food and NonFood prices indices: Compare scanner price indices (with and without quality adjustment) to BLS CPI
- NPD covers general merchandise and online retailers
  - NPD data have rich product attributes
  - Explore hedonics vs. alternative methods (e.g., UPI) for quality adjustment

Growth Rates of Survey vs. Scanner Data of Sales Track Each Other Well: Food



Price indices adjusted for quality at scale – Using same source data to measure revenue

Key challenge/opportunity: Enormous Product Turnover

- 650,000 products per quarter from 35,000 stores
- Product entry and exit rates (quarterly)
  - 9.62% (entry) and 9.57% (exit)
- Sales-weighted entry and exit rates
  - 1.5% (entry) and 0.3% (exit)
  - Rates vary substantially across product groups
  - Asymmetry in sales-weighted: “slow death” of exiting products
- Some of this entry/exit is substantive, other is marketing/packaging

Source: Nielsen scanner data (Food and NonFood)

Capturing product quality at scale: Alternative approaches

UPI: Expenditure function approach using CES aggregators (Redding and Weinstein, 2018, 2019)

- Capture product turnover with changing expenditure shares of new vs. old goods PV_{adj} (Feenstra 1994)
- Extend to capturing quality/appeal change of existing goods CV_{adj}
  - Captures ALL of the demand residual for quality adjusted prices.
- Needs item classification/nesting (all goods within a nest have equal substitutability)
  - How to do at scale?
- Requires estimation of elasticity of substitution for each nest.
- Requires defining common goods, entering and exiting goods
  - More complex than at first glance

Capturing product quality at scale: Alternative approaches

Hedonic approach with transactions data (Bajari and Benkard 2005, Erickson and Pakes 2011, Bajari et al. 2019)

- Estimate hedonic function within product groups using relationship between P and attributes on period by period basis
  - Use predicted hedonic prices for entering and exiting goods
  - Use chain weighting to continuously update weights
  - Both of these approaches helps accommodate product turnover
- For 21st Century Implementation, Need item attributes at scale
  - Bajari et al. (2019) provide guidance about machine learning methods that can be used at scale to:
    - Identify attributes from text and images
    - Use sophisticated nonlinear estimators to capture the relevant variation in a parsimonious manner for hedonic estimation.

Unified Price Index (UPI) (Redding and Weinstein 2018, 2019)

- Start with CES preferences for a given product group:

$$U_t = \left[\sum_{k \in \Omega_t} (\varphi_{kt} C_{kt})^{\frac{\sigma-1}{\sigma}} \right]^{\frac{\sigma}{\sigma-1}}$$

- Implies unit expenditure function (exact price index):

$$P_t = \left[\sum_{k \in \Omega_t} \left(\frac{p_{kt}}{\varphi_{kt}} \right)^{1-\sigma} \right]^{\frac{1}{1-\sigma}}$$

- φ_{kt} are time-varying appeal parameters
    - Normalization that keeps "average preferences" from shifting over time
- σ is elasticity of substitution, Ω_t are all goods in t
- Applied to narrow product groups (e.g. "Soft Drinks" or "Video Games")
    - Assume Cobb-Douglas utility over product groups – may have nests within product groups

Adjusting for quality via UPI

CES Demand function:

$$s_{kt} = \frac{p_{kt}c_{kt}}{\sum_l p_{lt}c_{lt}} = \frac{(\frac{p_{kt}}{\varphi_{kt}})^{1-\sigma}}{\sum_{l\in\Omega_t}(\frac{p_{lt}}{\varphi_{lt}})^{1-\sigma}} = \frac{(\frac{p_{kt}}{\varphi_{kt}})^{1-\sigma}}{P_t^{1-\sigma}}, k \in \Omega_t \qquad \text{All Goods}$$

s_{kt}^* is the expenditure share for common goods in period t-1 and t,

$\Delta ln\bar{s}_{kt}^* = \beta_0 + \beta_1\Delta\bar{p}_{kt} + u_{kt,}$ Doubled differenced equation with $\beta_1 = (1-\sigma)$.

Can estimate via Feenstra (1994) with assumptions about correlation of double differenced demand and supply shocks along with heteroskedasticity

With estimate of $\boldsymbol{\sigma}$ can recover quality factors. Need to keep track of product turnover and changing expenditure shares of common goods.
Critical issues: **ALL of demand shock included in product quality. Level of Aggregation**

Unified Price Index (UPI) (Redding and Weinstein 2018, 2019)

$$\log(\text{UPI}) = RPI + PV_{adj} + CV_{adj}$$

RPI is Jevons Index

$$RPI = \frac{1}{N_t^*} \sum_{k \in \Omega_t^*} \ln\left(\frac{p_{kt}}{p_{kt-1}}\right) \qquad PV_{adj} = \frac{1}{\sigma - 1} \ln\left(\frac{\lambda_t}{\lambda_{t-1}}\right) \qquad CV_{adj} = \frac{1}{\sigma - 1} \frac{1}{N_t^*} \sum_{k \in \Omega_t^*} \ln\left(\frac{s_{kt}^*}{s_{kt-1}^*}\right)$$

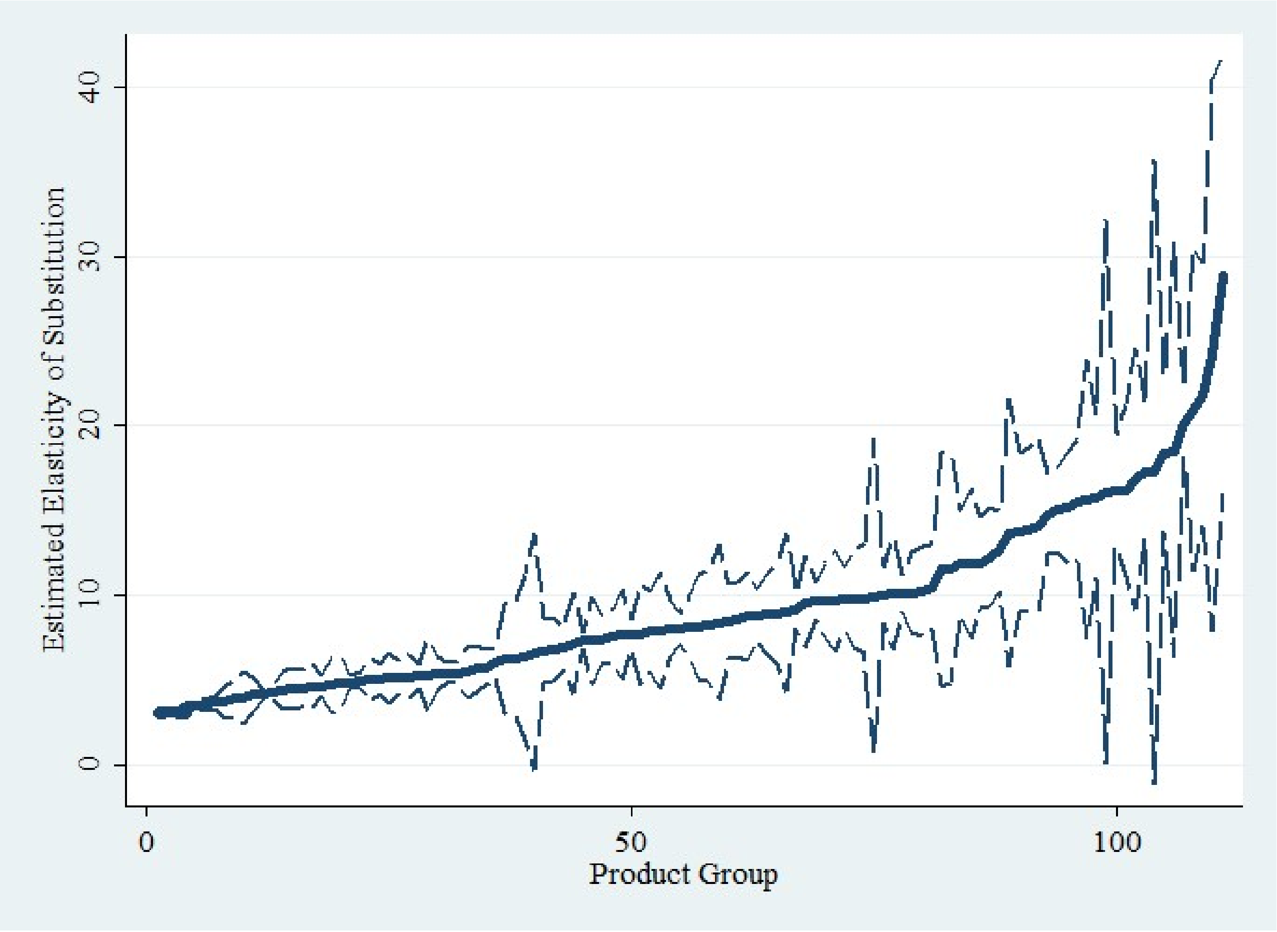
Key Issues:

- Magnitude of adjustment factors depend on elasticity of substitution for narrow group.
- $\lambda_t \equiv \frac{\sum_{k \in \Omega_t^*} P_{kt} C_{kt}}{\sum_{k \in \Omega_t} P_{kt} C_{kt}}$ where Ω_t^* are common goods and Ω_t are all goods in t.
- s_{kt}^* is volatile for recently entered and goods about to exit. CV_{adj} sensitive

to definition of **COMMON GOODS**.

- Intermediate step: Only consider product turnover yields “Feenstra” index:
  $\log(Feenstra) = PV_{adj} + SV$, where SV is the Sato-Vartia index

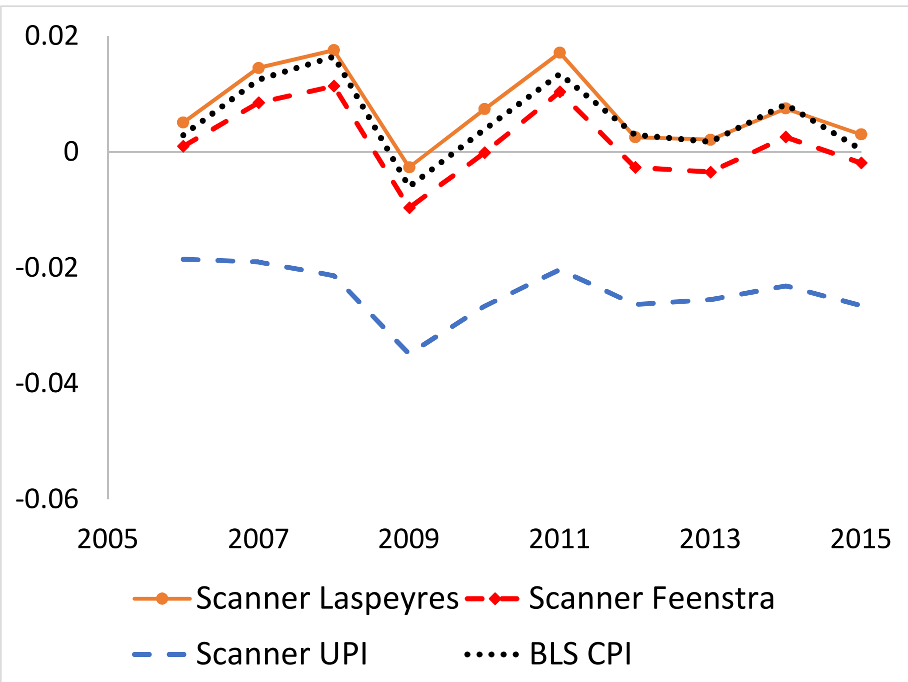
Estimated Elasticities of Substitution by Product Group



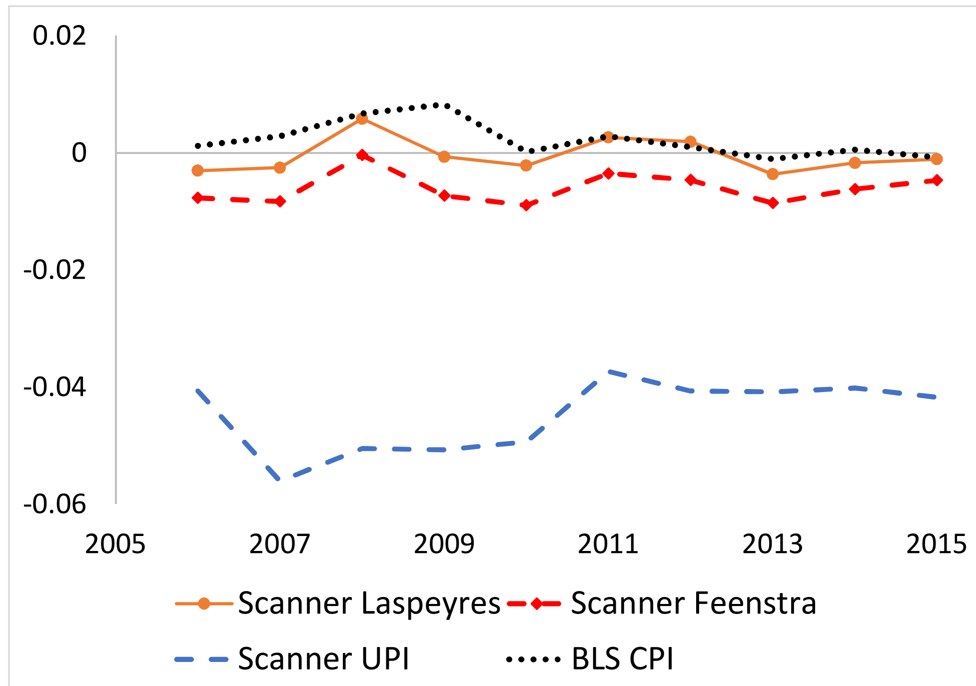
Laspeyres index using scanner is similar to BLS CPI (especially food), Feenstra somewhat lower and UPI much lower.

This implementation of UPI at quarterly frequency where COMMON goods are those present in t-1 and t.
RW (2019) show that CV_{adj} is reduced substantially when COMMON goods are defined over LONGER HORIZON.

Food



Nonfood



In practice, also some sensitivity to using Nielsen Consumer Panel vs. Nielsen Scanner. Likely related to small S_{kt}^*

Hedonics and transactions data

Following Bajari and Benkard (2005) and Erickson and Pakes (2011) hedonics regressions estimated every period using item-level data

$$p_{it} = X_i'\beta_t + \eta_{it}, \text{ where } X_i \text{ is vector of characteristics}$$

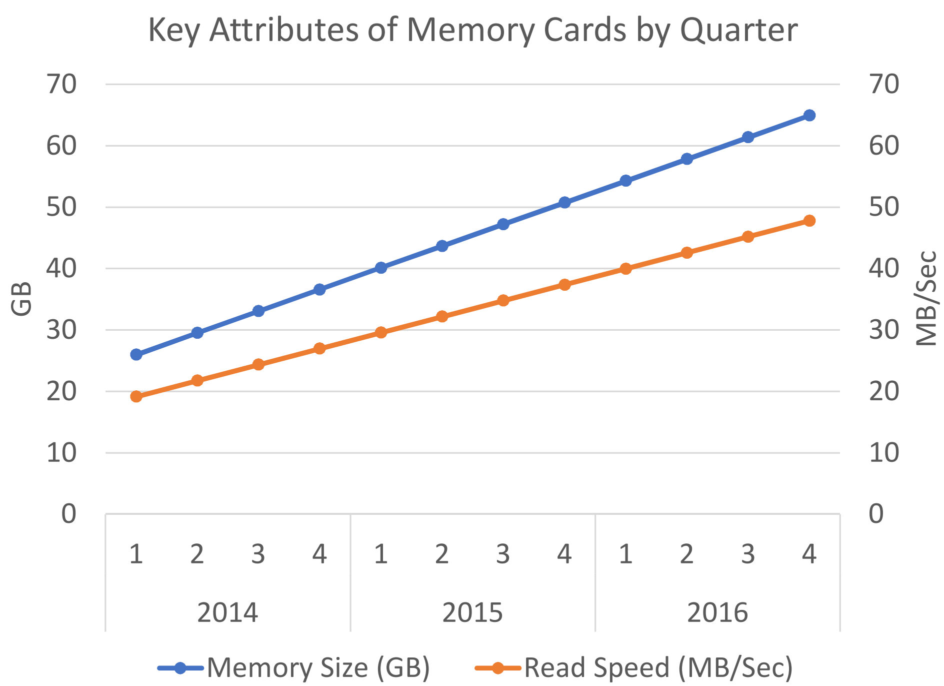
Laspeyres Hedonic Index (*LPH*) given by

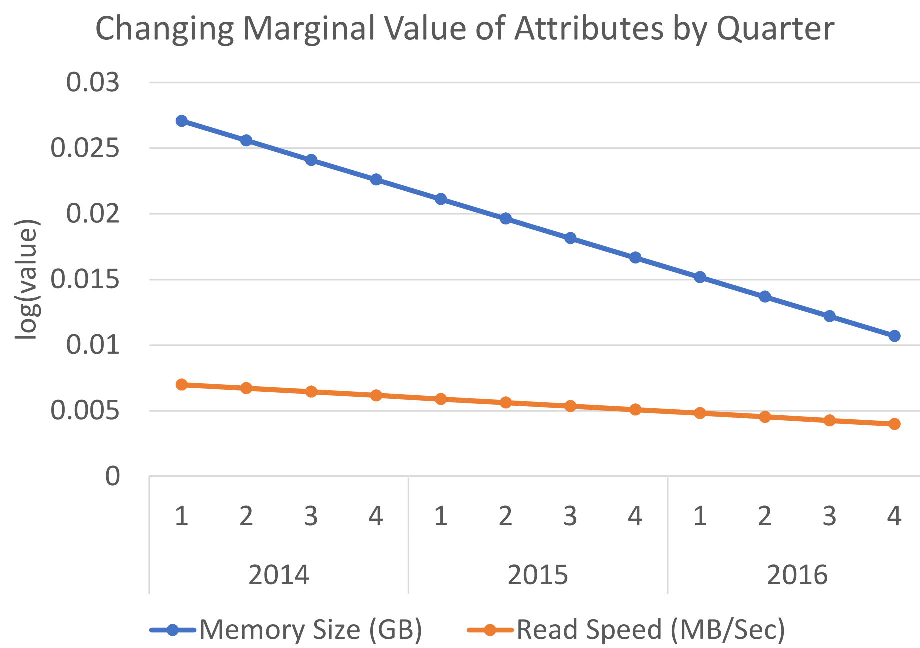
$$LPH_t = \frac{\sum_{i \epsilon A_{it-1}} h^t(X_i) q_{it-1}}{\sum_{i \epsilon A_{it-1}} h^{t-1}(X_i) q_{it-1}}$$

where $h^t(X_i)$ is the period *t* estimate of the hedonic function and A_{it-1} is the set of all goods sold in period t-1 (including exits). Use predicted hedonic prices for entering/exiting goods.

Critical issues: Requires measuring characteristics. Omitted unobserved/unmeasured characteristics cause biases.

NPD item-level characteristics for Memory Cards
Quality improves over period; marginal value falls

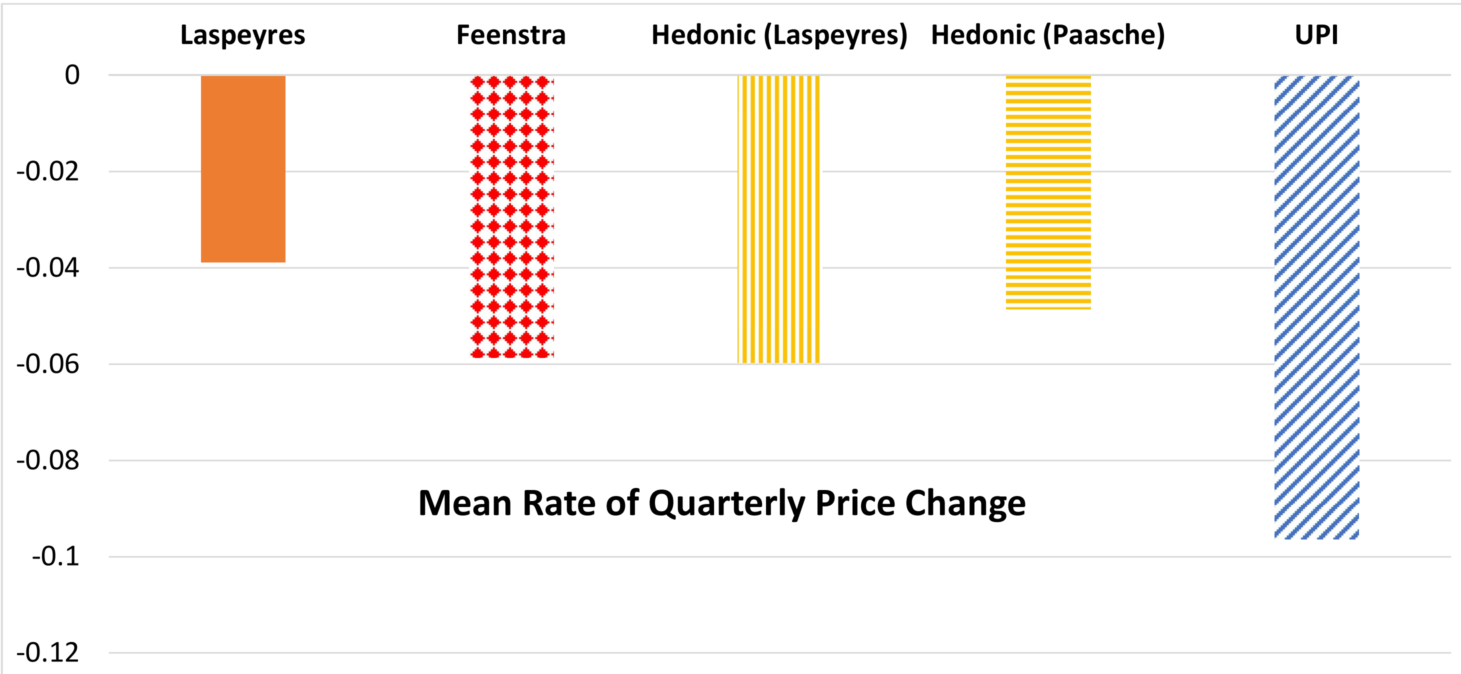




Linear trend on sales-weighted Memory size and speed.

Trend of linear terms from hedonic regression

Alternative Price Indices Memory Cards



Correlations with UPI	Laspeyres	Feenstra	Hedonic (Laspeyres)	Hedonic (Paasche)
	0.15	**0.07**	**0.32**	**0.48**

Key attributes for Memory Cards: Size and Speed, R-squared for Hedonics is about 0.8 each quarter
Elasticity of Substitution is 4.21

Open Questions:

- **In principle, both UPI and hedonic approaches can be done at scale using item-level transactions data.**
  - **Both approaches accommodate product turnover and quality adjustment**
- More research is needed:
  - What is the relationship between quality adjusted price indices using these two distinct approaches?
  - Are the UPI and hedonics likely to “converge” if UPI is based on nests defined by comprehensive attributes used to estimate hedonics?
  - What do we learn about conceptual and measurement issues by examining the differences in the price indices generated from the alternative approaches?
  - Can we use the alternative approaches for cross validation?