

Measuring Retail Trade with Administrative Data: U.S. Bureau of Labor Statistics

David Friedman

Associate Commissioner for Prices & Living
Conditions

Federal Economic Statistics Advisory Committee

June 10, 2016

Data Sources

- Administrative/Publicly available data
- Purchased data sets
- Company provided data – “corporate level data”
- Web scraping/ application program interface (API)



CPI Data Uses

- Create sample frames
- Benchmark samples
- Supplement collected data to support hedonic modeling (quality adjustment)
- Replace/supplement current data collection methods



Summary: Replacing Collection Initiatives

- Almost complete

- ▶ CorpY – company provided dataset

- In progress

- ▶ CorpX – company provided dataset

- ▶ JD Power – purchased data

- ▶ Nielsen – purchased data

Corporate Level Data: CorpY

➤ Great Opportunity

- ✓ maintain respondent cooperation
- ✓ reduce respondent burden
- ✓ work with transaction level data
- ✓ receive insurance prices

➤ Challenges

- ✓ Average prices for broader category and aggregated
- ✓ Data received in format difficult to process
- ✓ Status: 1st production use is May 2016 Index for monthly quotes

Corporate Level Data: CorpX

- Receive sales data monthly by 5th of following month
- Great Opportunity
 - ✓ maintain respondent cooperation
 - ✓ reduce burden
 - ✓ work with sales data



Corporate Level Data: CorpX

➤ Challenges

- ✓ mapping the CorpX item categories to the CPI structure
- ✓ melding the sales level data into our methodology and current system
 - in particular, accommodate seasonality & item substitution including new methodology
 - achieve constant-quality price change w/a big data set
- ✓ lack of characteristic detail
- ✓ having enough history to validate method

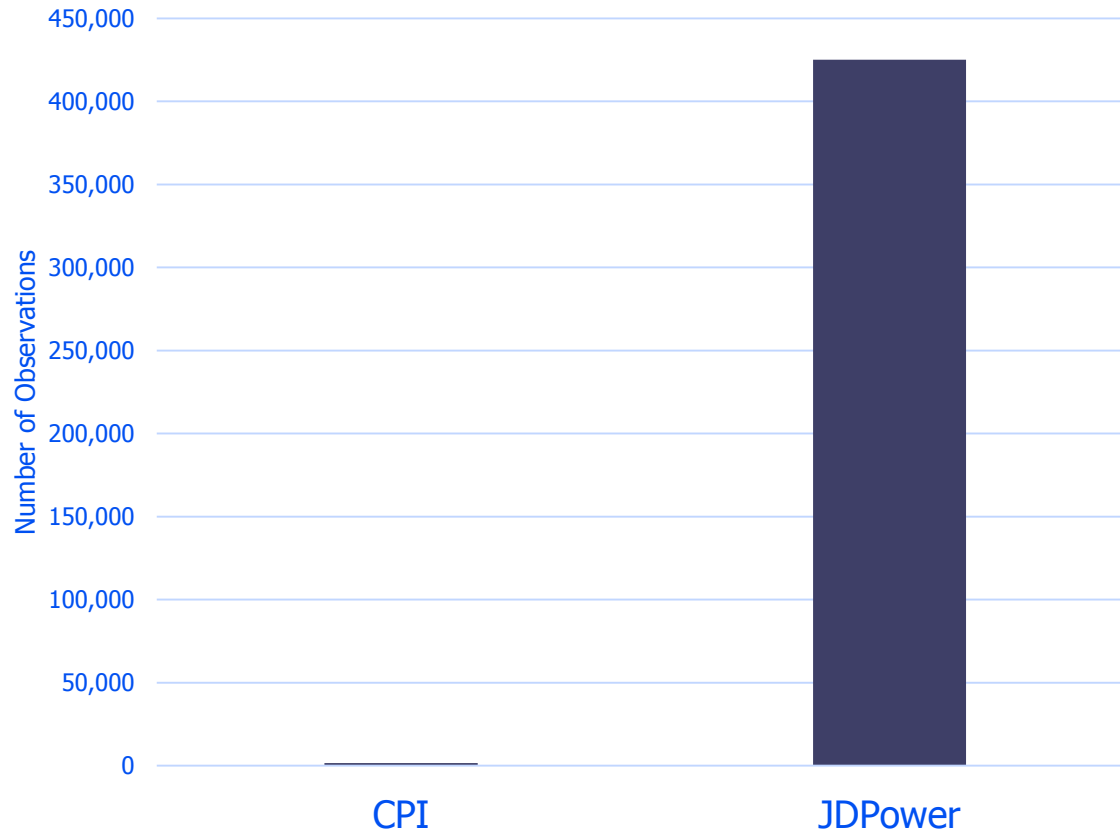
CorpX Current Status

- I. Received data for all CPI Primary Sampling Units (PSU's) beginning with October 2014
- II. Testing various methodologies
- III. Will develop necessary CPI system changes to be ready to use

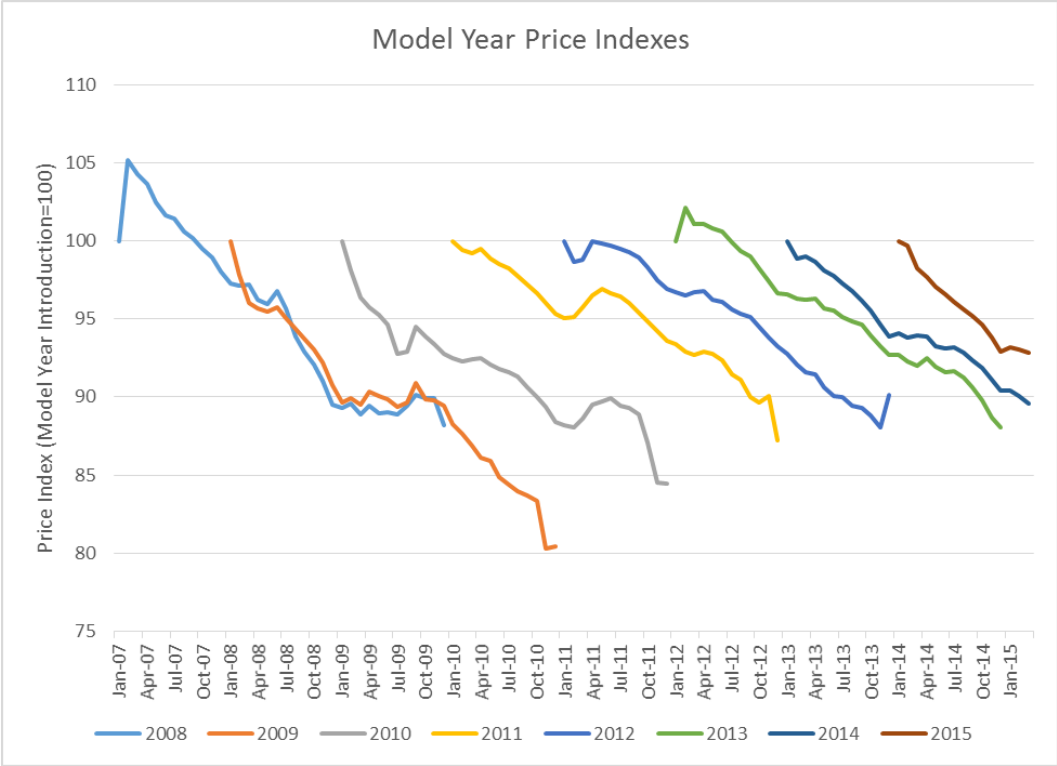
JD Power Project

- Purchase JD Power dataset as source for replacement in New Vehicles index
- Prime example of benefits and challenges of “big data”
 - ▶ Breadth of information
 - ▶ Challenge of integration with current systems
 - ▶ Methodological issues

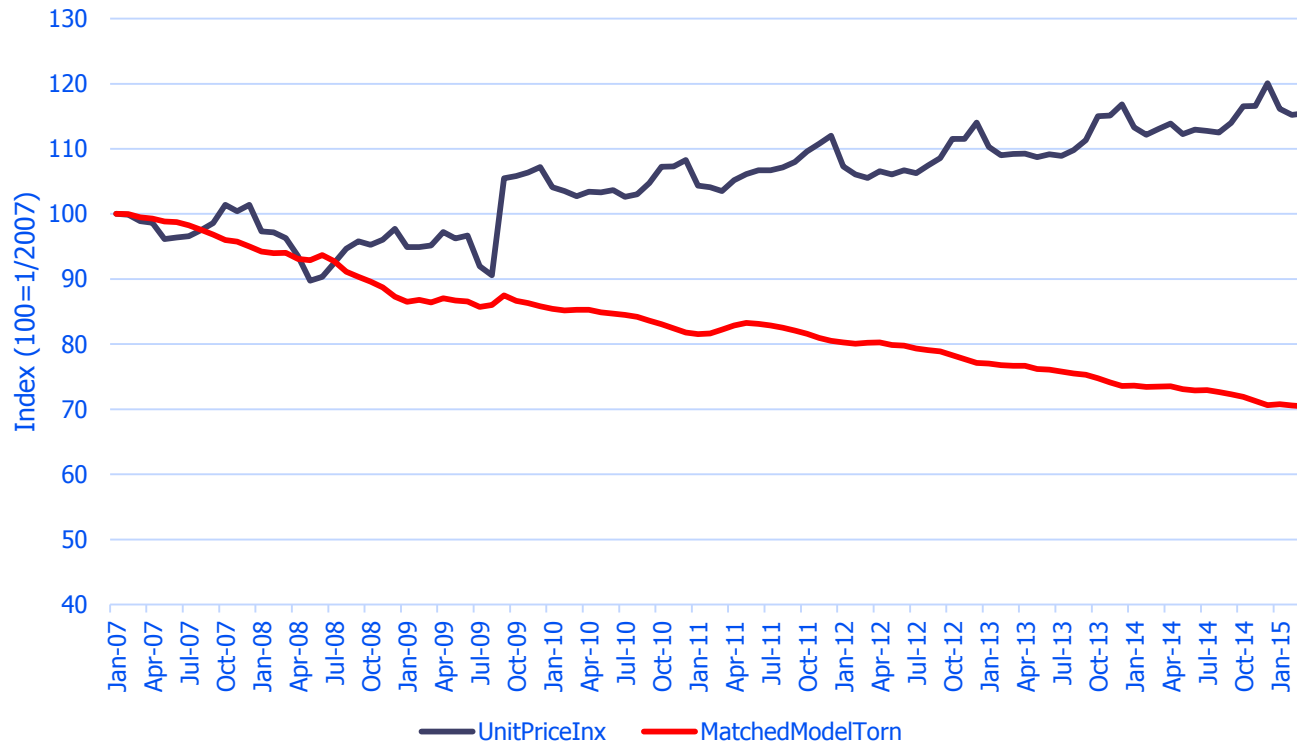
New Vehicle Observations



Model Year Price Indexes



Unit Prices Increase



Ways to Treat the Price Declines

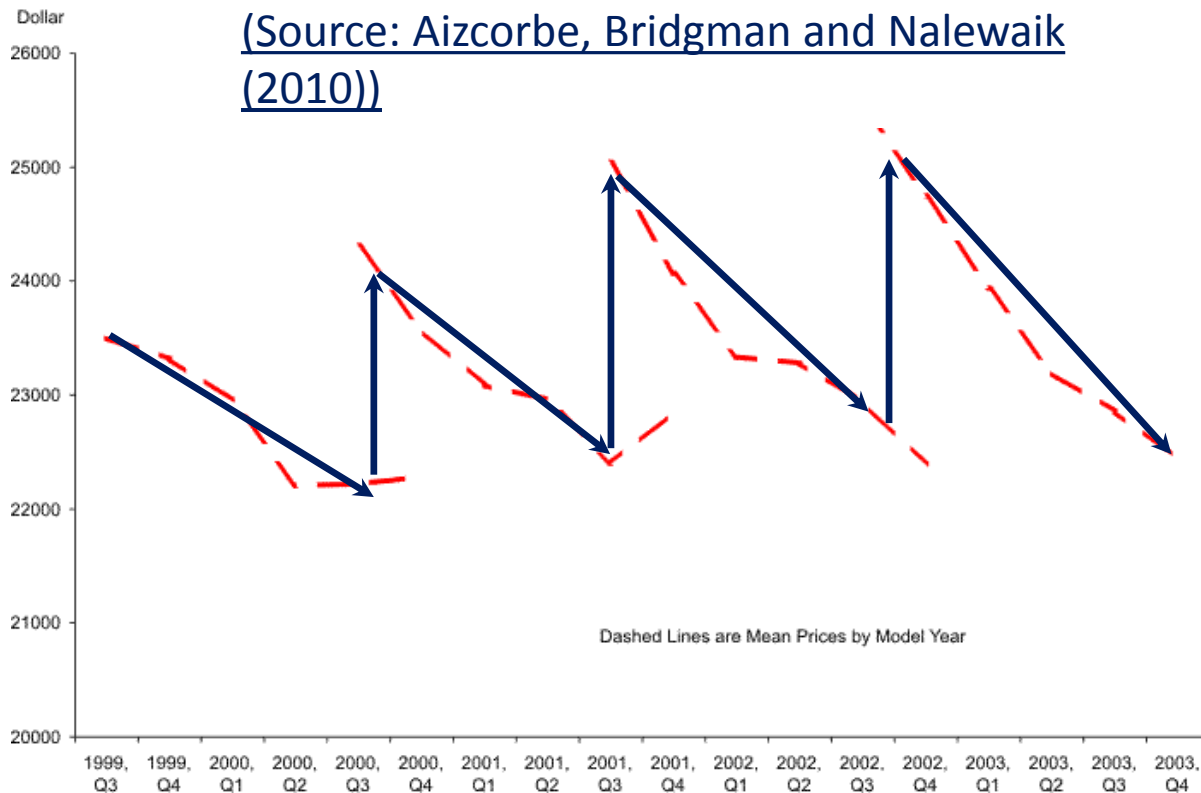
- Show the drop
- Show price change across model years
 - ▶ Create “Changeover” price relatives
 - ▶ Use Year-Over-Year Index



Price Dynamics

Average Prices

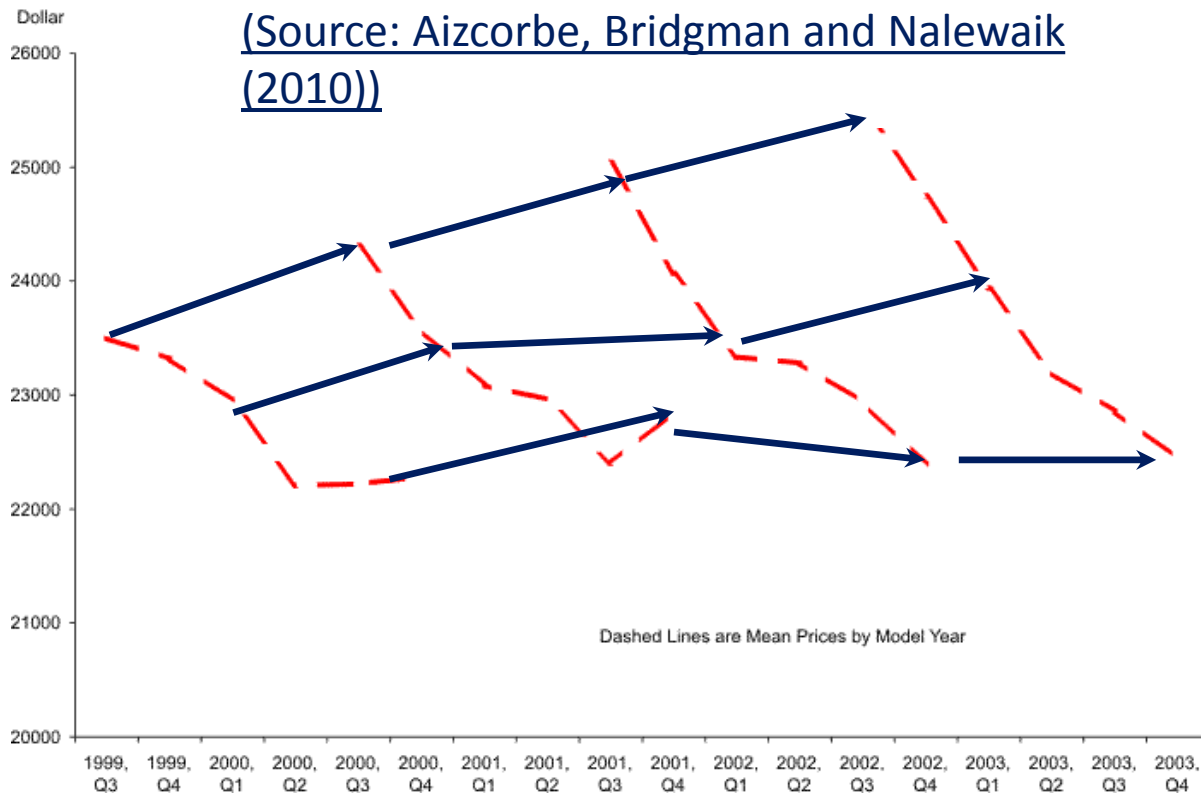
(Source: Aizcorbe, Bridgman and Nalewaik (2010))



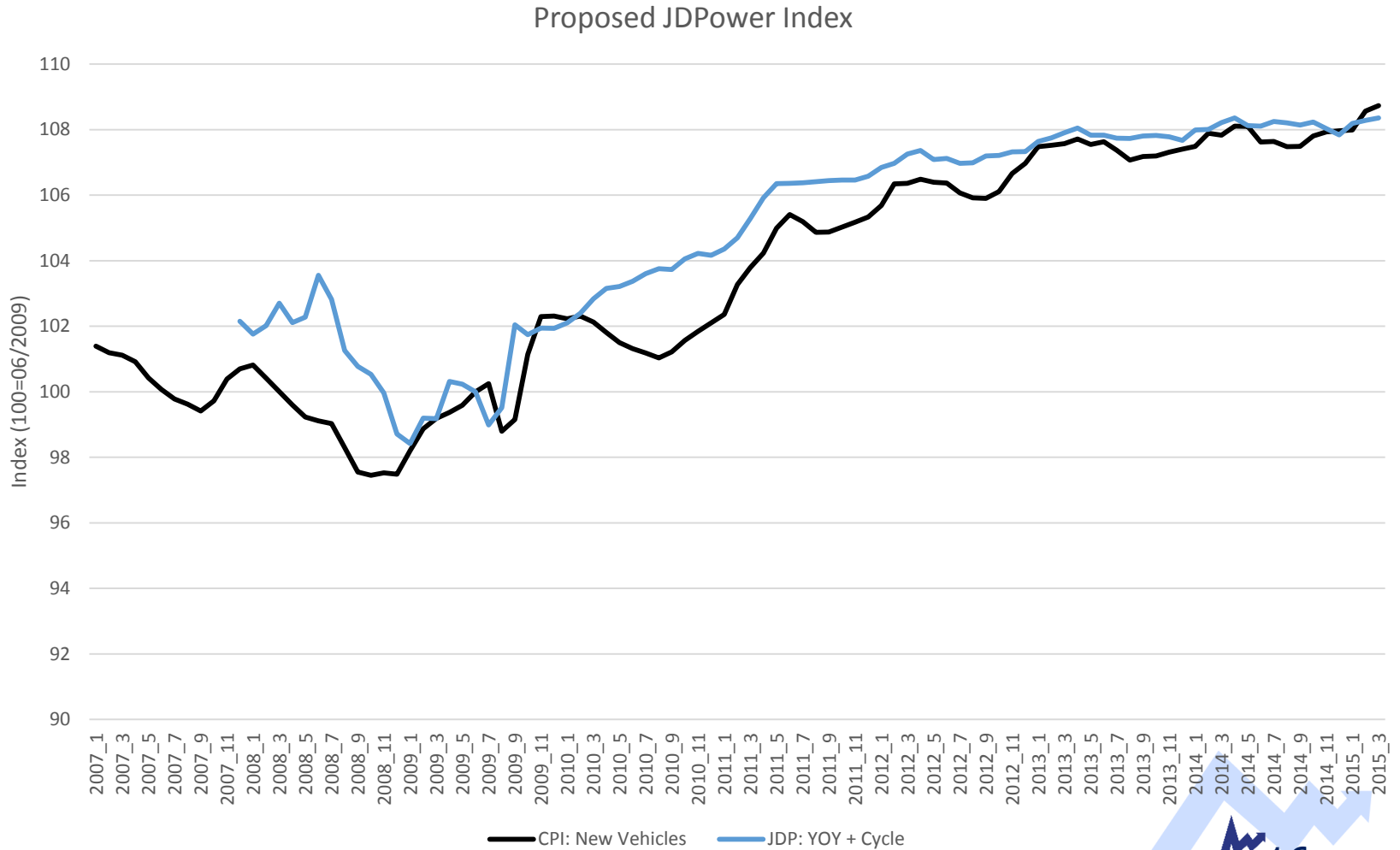
Price Dynamics

Average Prices

(Source: Aizcorbe, Bridgman and Nalewaik (2010))



JDPower vs CPI

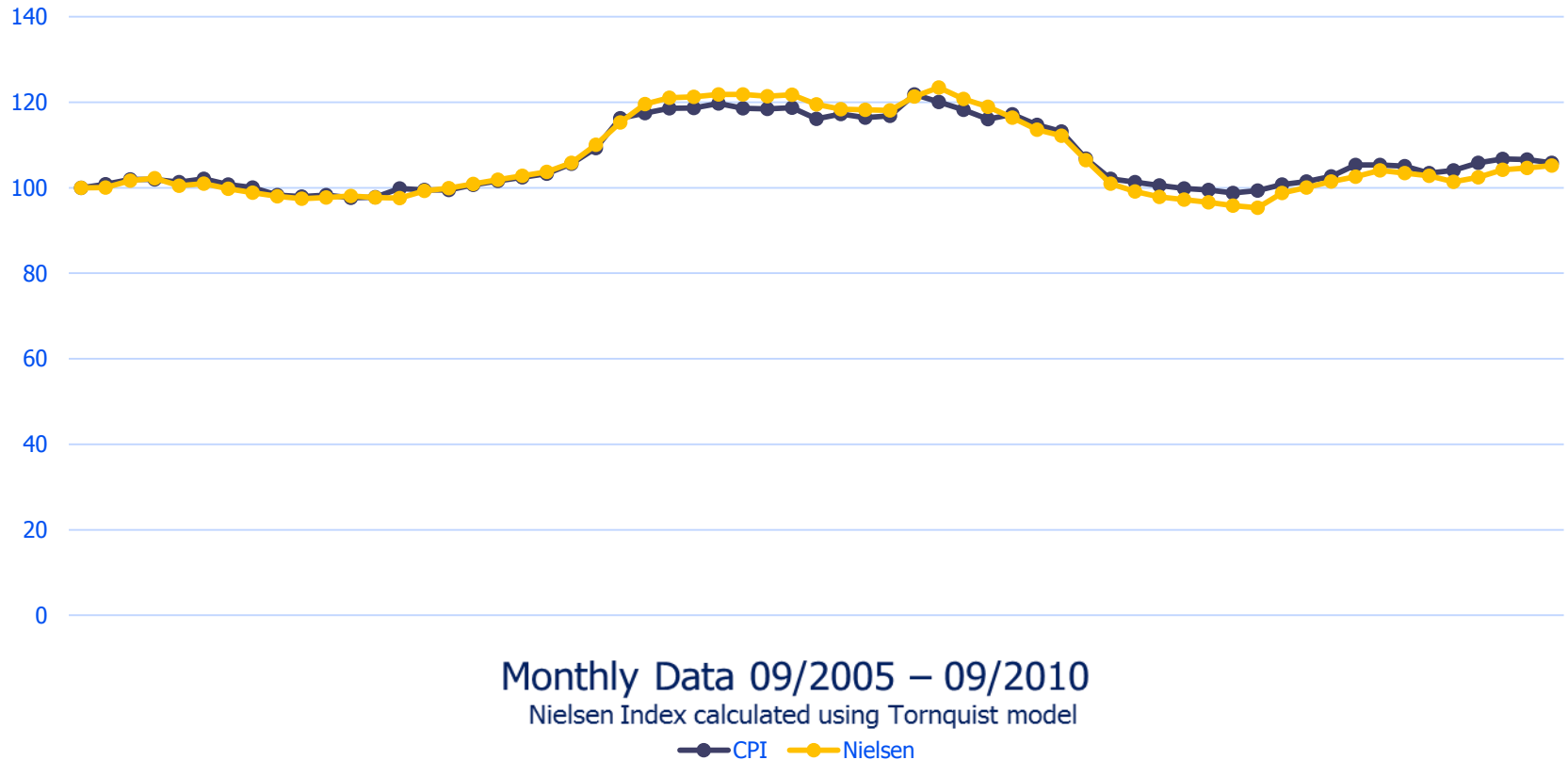


Research Nielsen Indexes

- Data set for August 2005 – September 2010
- 2 million UPC codes
- Scantrack coverage limitations
 - ▶ Grocery > \$2 million; Drug Stores > \$1 million; Mass Merchandisers
 - ▶ Excludes one major retailer and non-UPC items (some produce, deli, bakery, fresh meat, etc.)

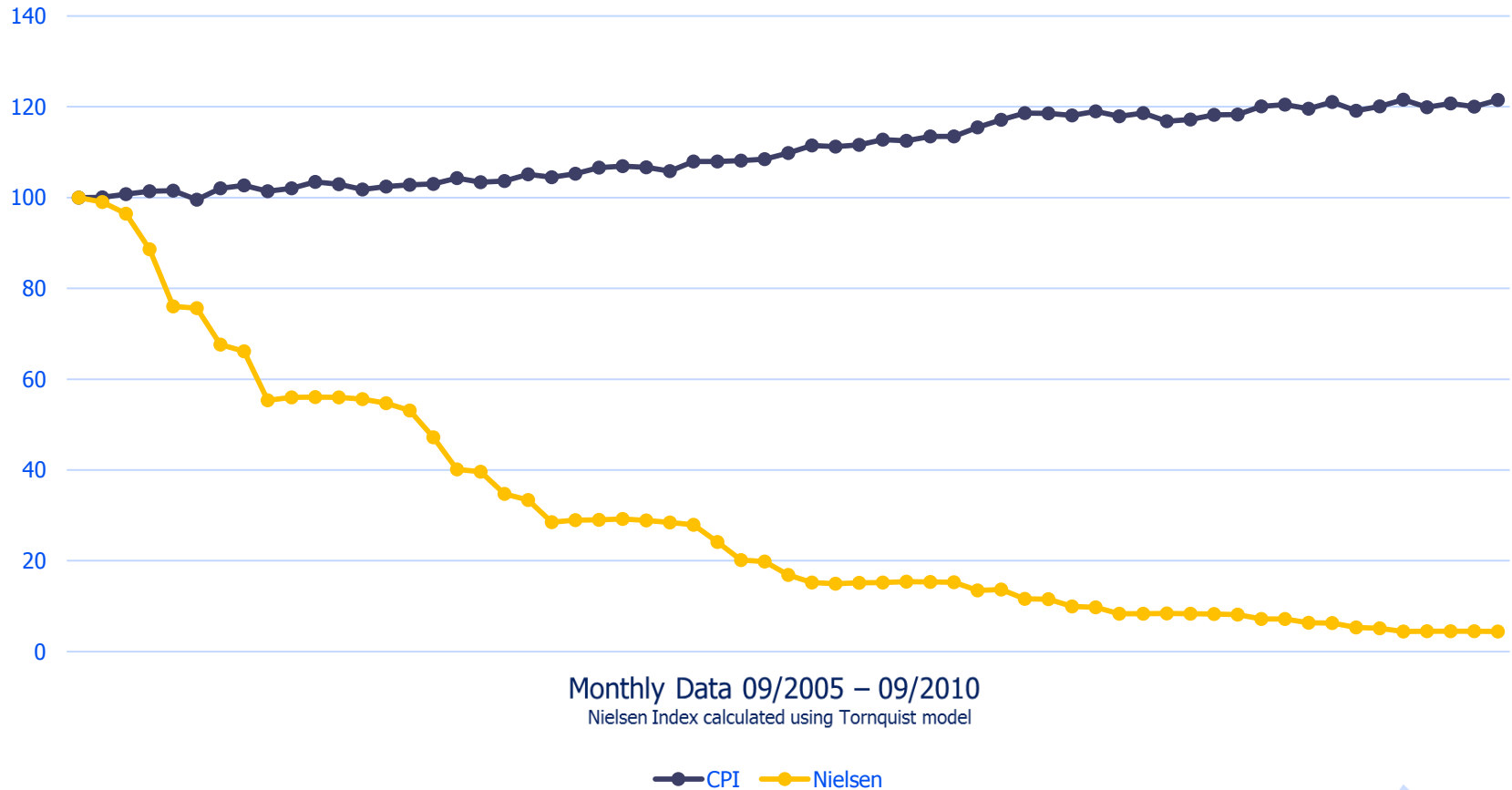
Nielsen Indexes

FJ011 - Milk



Nielsen Indexes

FR02 - Candy and chewing gum

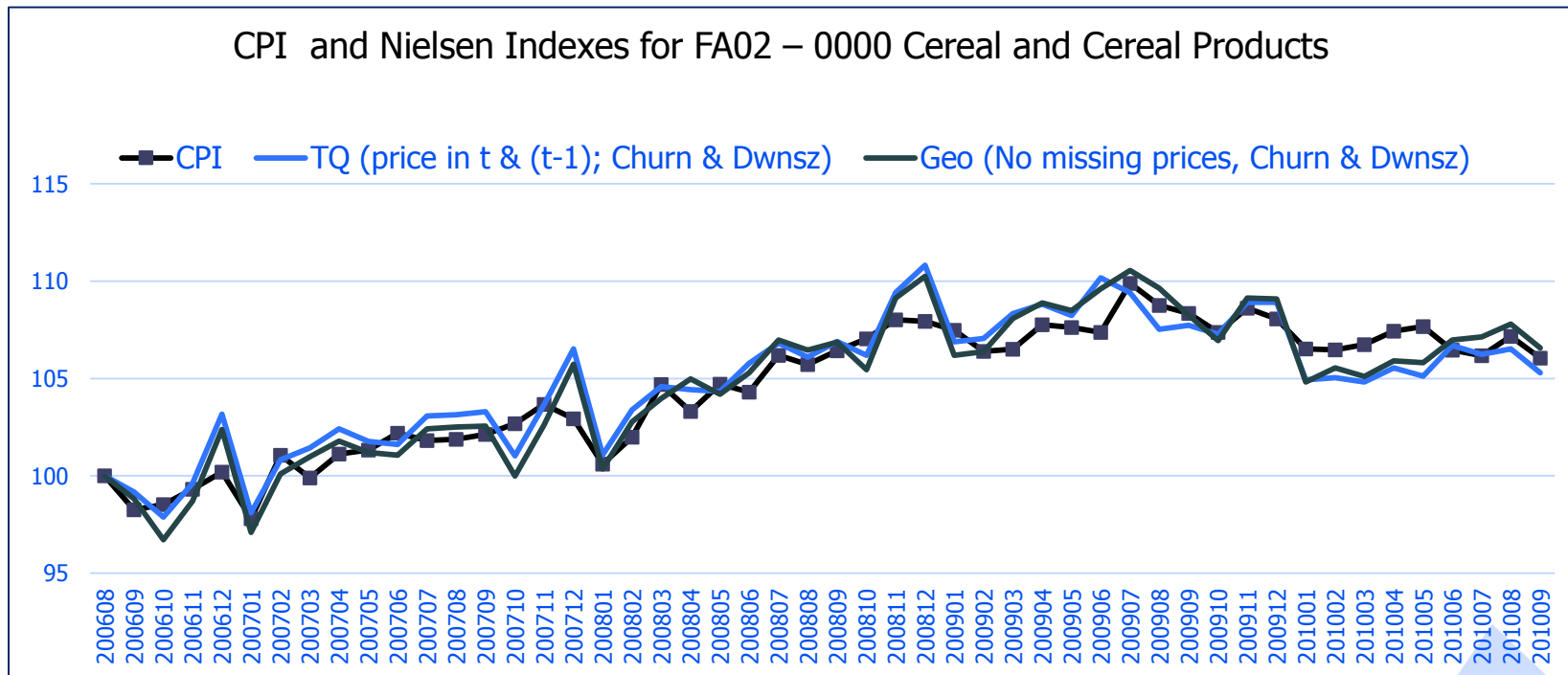


Nielsen Indexes – Current focus

- Refine Nielsen indexes to :
 - ▶ Limit research to items that are well represented in the Scantrack data
 - ▶ Account for product downsizing
 - ▶ Account for UPC “churn”
 - ▶ Calculate a geomeans index (in addition to a Tornqvist index)

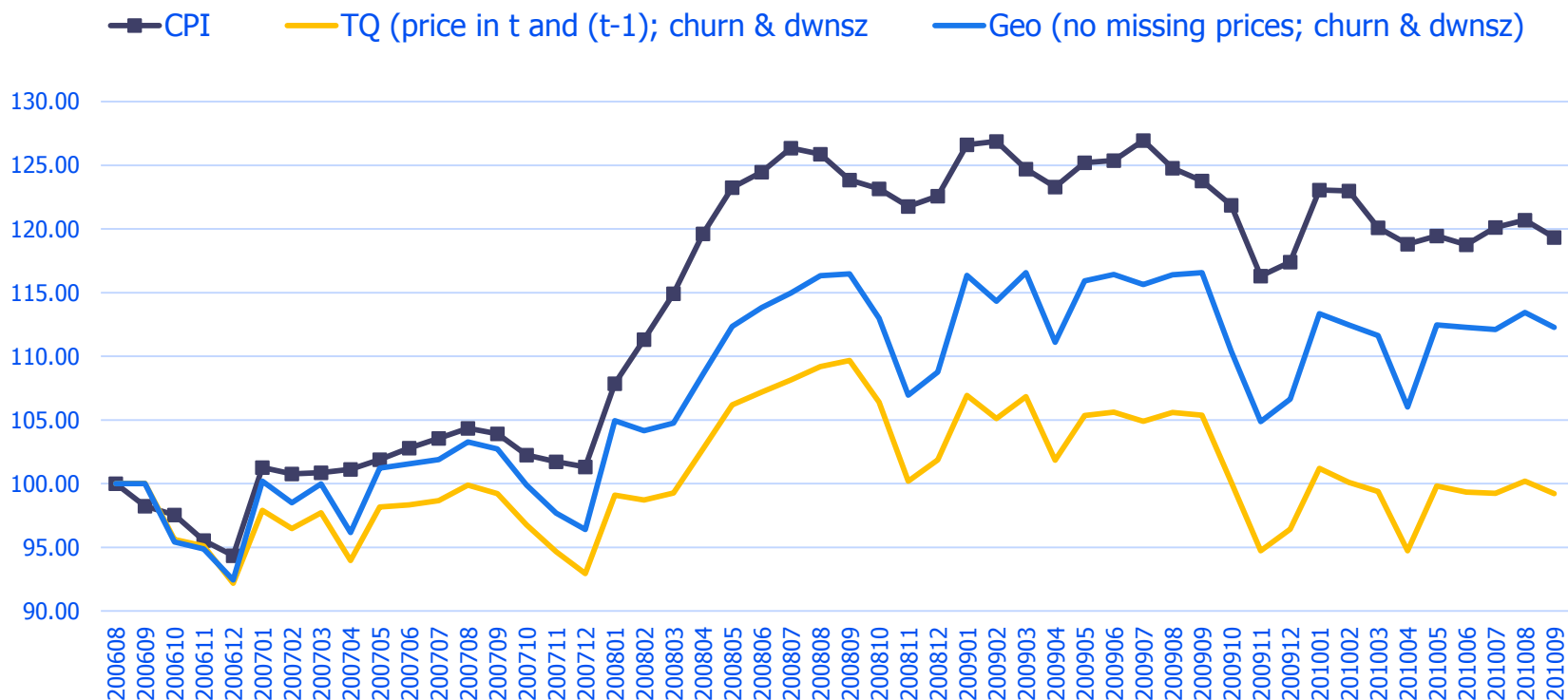
Nielsen indexes – Current focus

- Preliminary results for 4 item strata
- Work on additional 10-12 strata in FY16



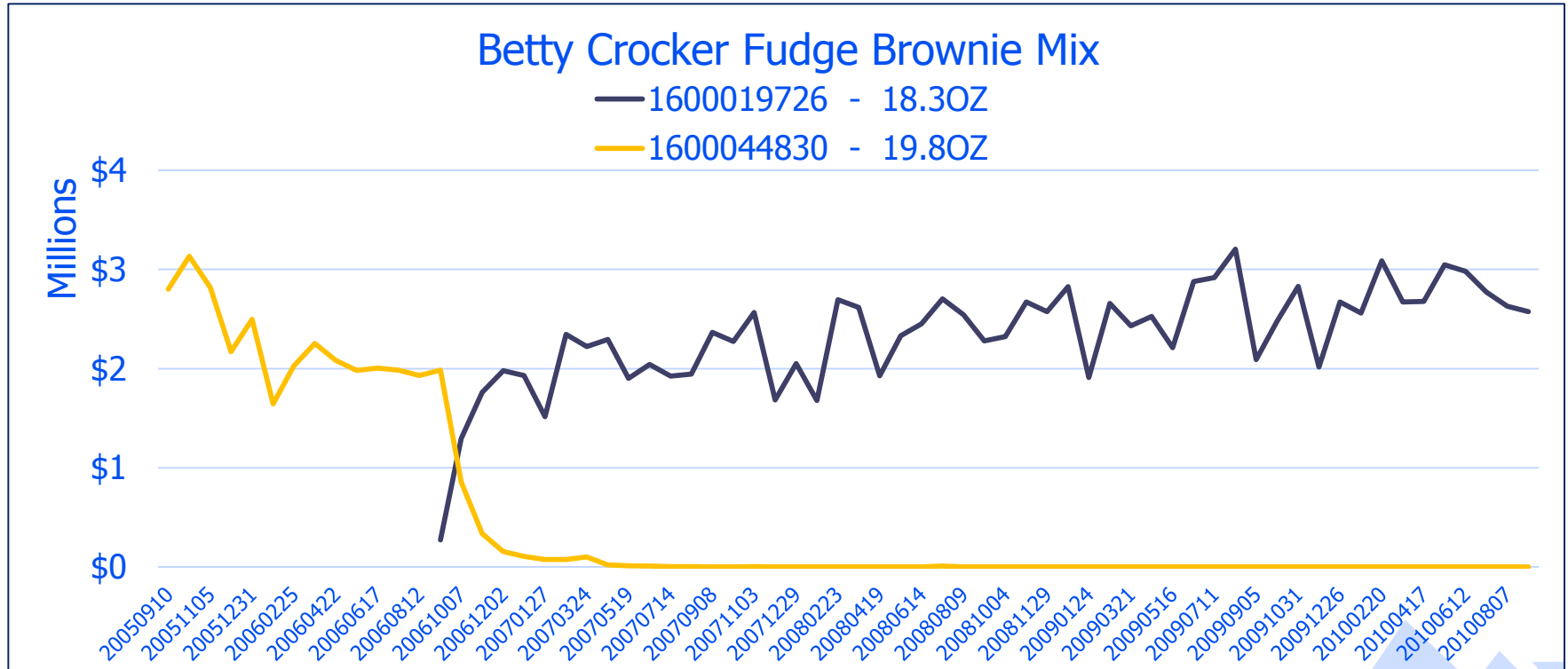
Nielsen Indexes – Current focus

CPI and Nielsen Indexes for FA01 – 0000
Flour & Prepared Flour Mixes



Nielsen downsizing

- Automate identification
- Compare to CPI



Summary: Benefits vs. Challenges

Benefits:

- Increasingly more available
- Allows for evaluation & improvement
- May reduce collection costs
- Reduces respondent burden
- Increased sample size
- May increase data quality
- Sometimes ability to get quantity data

Challenges:

- Data quality issues – especially lack of descriptive info
- Timeliness and reliability concerns – mitigation strategies
- Cost and other considerations (new skill set, IT infrastructure, etc.)

What's Next

- Continue work on CorpX, JD Power, Nielsen
- Project to modify CPI production to more readily accept future alternative data
- Work with CE to investigate secondary sources for Rent Data
- Explore new opportunities



Contact Information

David Friedman

Associate Commissioner for Prices & Living
Conditions

www.bls.gov/bls/inflation.htm

202-691-6307

Friedman.David@bls.gov

