RESET Project

"Re-engineering Official Statistics through Big Data and AI"

Matthew D. Shapiro

University of Michigan and NBER

ESRI Policy Forum

Cabinet Office, Government of Japan

Tokyo – 30 July 2025

Good Policy Requires Good Data: A Parable

Policy debate in January 2009

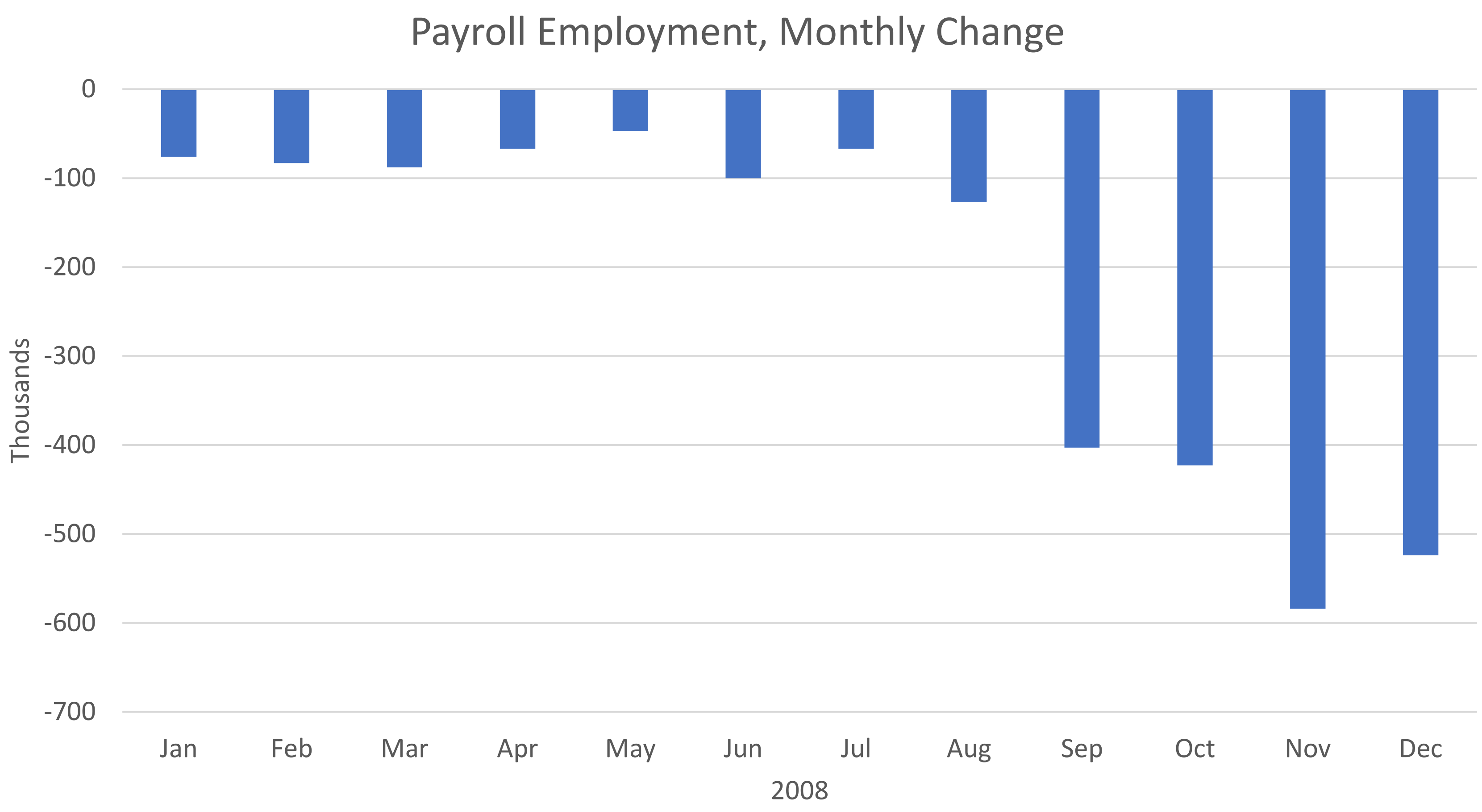
Questions:

- How bad was the economy?
- How big a stimulus package?

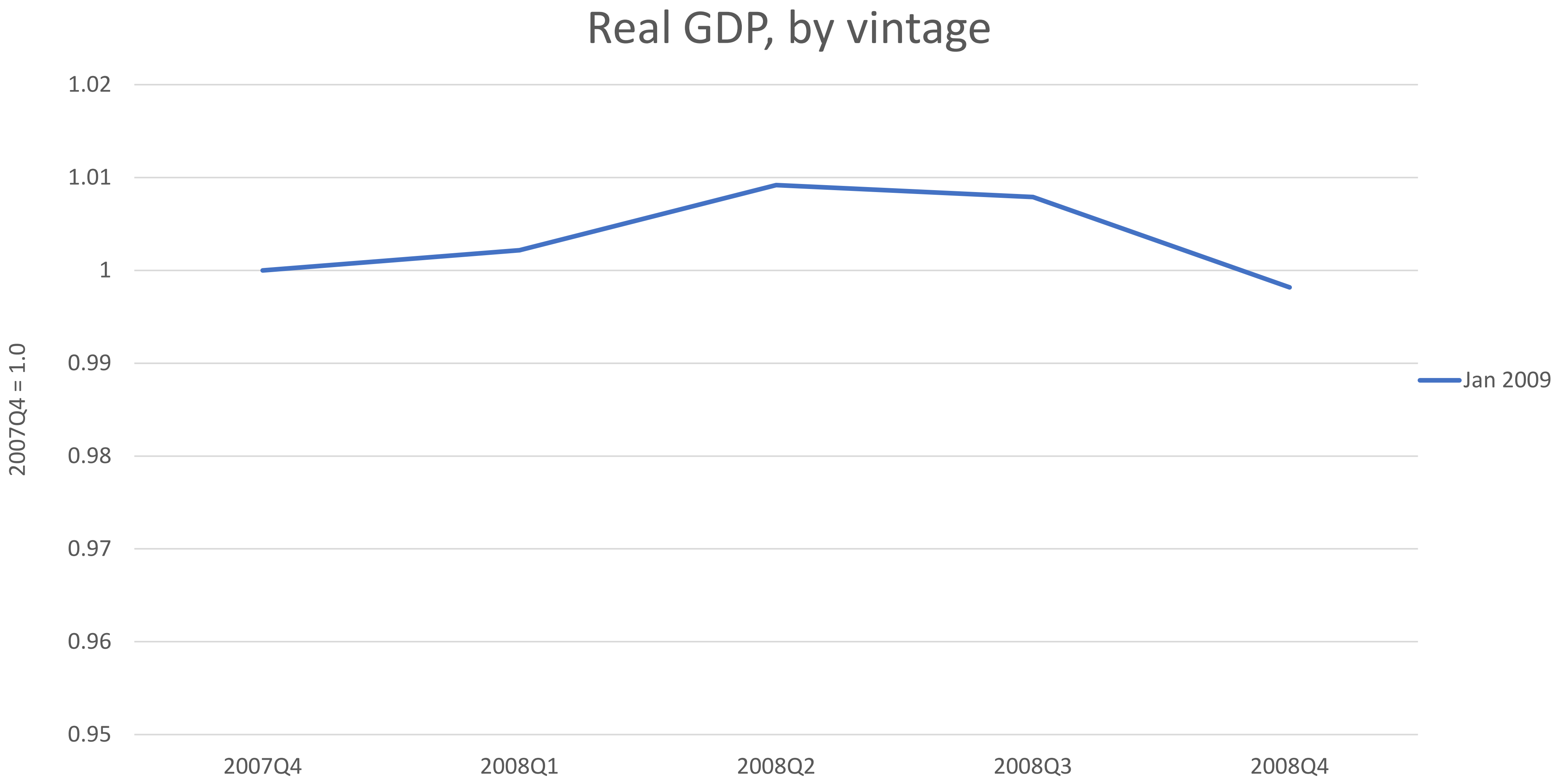
Policy perspectives:

- Obama administration: Significant stimulus warranted
- Congress: Deficit concerns

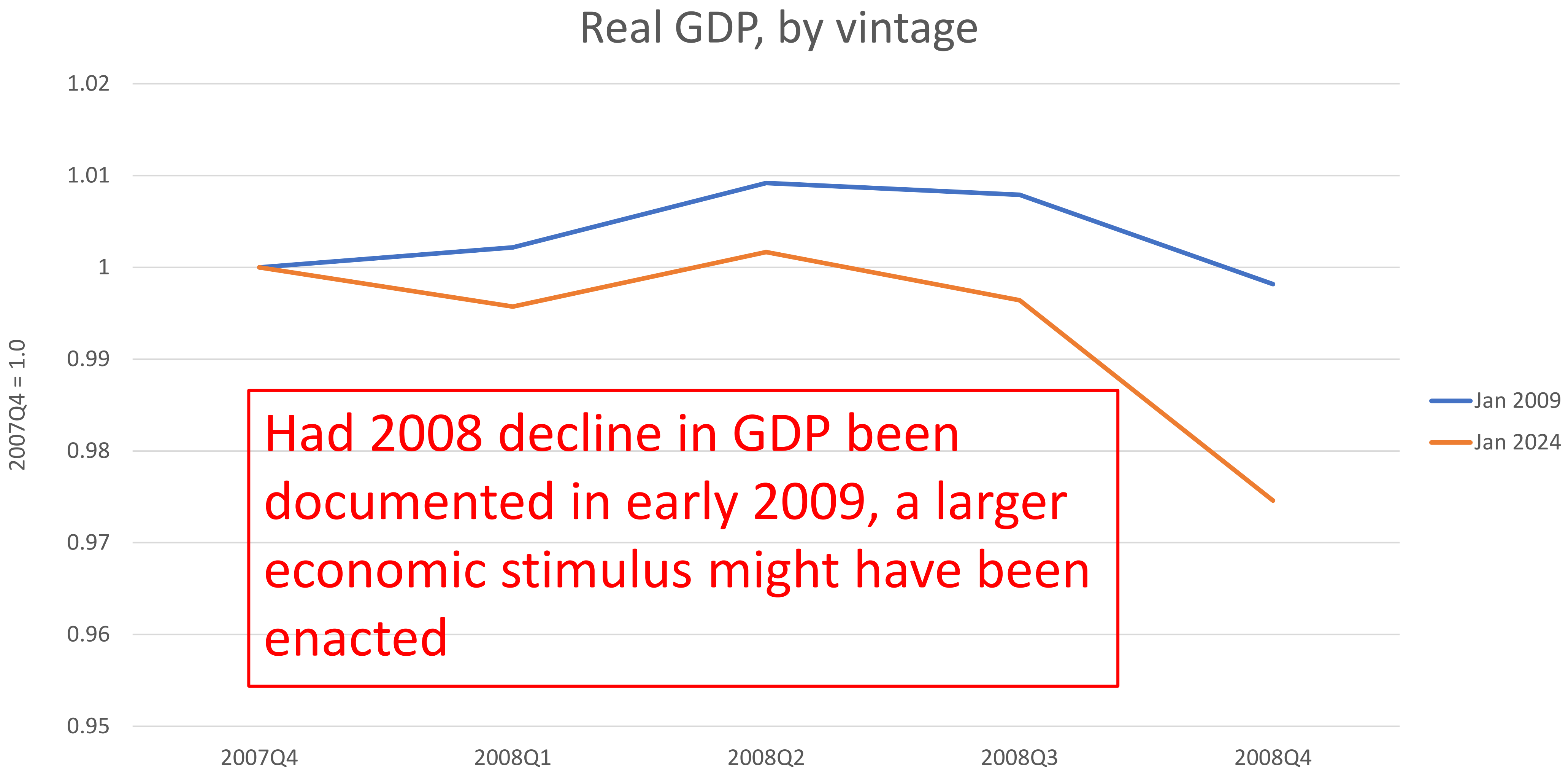
How bad was the economy in late 2008? Employment was plummeting



But real GDP falling, but not dramatically



But real GDP falling, but not dramatically—Until revisions!



Multiple sources of revisions

- Long lags in availability of data
  - Benchmark business data to quinquennial economic censuses
  - High-frequency data very aggregated (e.g., store-level)
  - Tax administration data lags (filings, subsequent processing)
- Limited use of real-time business data (transactions, payrolls)
- Extrapolation and interpolation for preliminary estimates

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

Re-engineering measuring sales and prices

Challenge: Tap the firehose of transactions level data
now available from businesses on P and Q .

Firms' item-level transactions data
- Brick and mortar
- Internet retailers
- Aggregators

Agencies
Data products:
- GDP
- inflation

Data improvements:
- Quality change
- Timeliness
- Granularity
- Distributional statistics

RESET Project:
Re-Engineering Statistics using Economic Transactions

Aims

- Measure consumer price inflation and components of GDP from item-level transactions (“scanner data”)
- 21st Century statistical infrastructure:

  Replace current system of surveys and enumerations with a pipeline from business information systems to statistical agencies

RESET Project:
Re-Engineering Statistics using Economic Transactions

Benefits for research/policy/economy/society: Improved data quality

- Internally consistent measures of inflation and real expenditure
- Timeliness
  - Near real time; high frequency
  - “Final” data available almost immediately; Avoids revisions inherent in current infrastructure
- Improved granularity (frequency, product/industry detail, geography)
- **Improved measurement inflation**
  - **Superlative price indexes require simultaneous measurement of P and Q**
  - **Measuring ubiquitous quality change: Use machine learning/AI to manage firehose**

RESET Project:
Re-Engineering Statistics using Economic Transactions

Benefits for firms

- Better data: Public good
- Better data: Private good, esp. having official statistics more consistent with internal firm data
- Potentially less respondent burden
    - Reduce/eliminate surveys (some of which are redundant)
    - Replace in-store or phone price enumerations
    - Use third-party reports (i.e., aggregators)

Re-Engineering statistics is not optional…

- Response rates to firm surveys declining
  - E.g., Advance Retail Trade Survey (backbone of monthly Personal Expenditure) has a response rate of two-thirds
  - Non-mandatory: Firms enter and exit survey at high frequency, complicating estimation
- BLS: steps toward scanner data
  - Some pharmacy and retail chains prefer sharing data to having enumerators in stores
  - Web-scraped, crowd-sourced data
- RESET principle: **Collect the quantities along with the prices**
  - Needed for **superlative price indexes**
  - **Nominal spending**
  - **Needs cooperation of firms: quantity sold cannot be scraped**
- RESET focuses on goods, but lots of services are scanned (airlines, hotels, utilities, rent, etc.)

RESET: Broad set of collaborations and data sources

- **Aggregators**
  - Circana [formerly NPD and IRI] via Census Bureau:
    - Broad range of general merchandise and consumer packaged goods
    - Circana partner in new RESET Demonstration Project
  - NielsenIQ via Chicago Kilts Centers:
    - Grocery, discount, and liquor stores and pharmacies
  - NielsenIQ data via Census Bureau
- **Collaborations with large retailers**
  - Working within data infrastructure of one large retailer
  - NDA stage with other retailers

RESET: Methodological challenge Measuring quality change at scale

Digitized Prices and Quantities Reveal Enormous Product Turnover

- High product (item-level) entry and exit rates
- Some turnover is substantive, some is marketing/packaging
- Traditional matched model price indices ignore this turnover

  → Need to address massive product turnover to get advantages of granular, real-time data

Air Jordan XXXVI Low
Men's Basketball Shoes
$175
Style: DH0833-063

Air Jordan XXXVII
Men's Basketball Shoes
$185
Style: DD6958-060

Air Jordan XXXVI versus XXXVII:
<u>Does the $10 price difference owe to higher quality</u>?

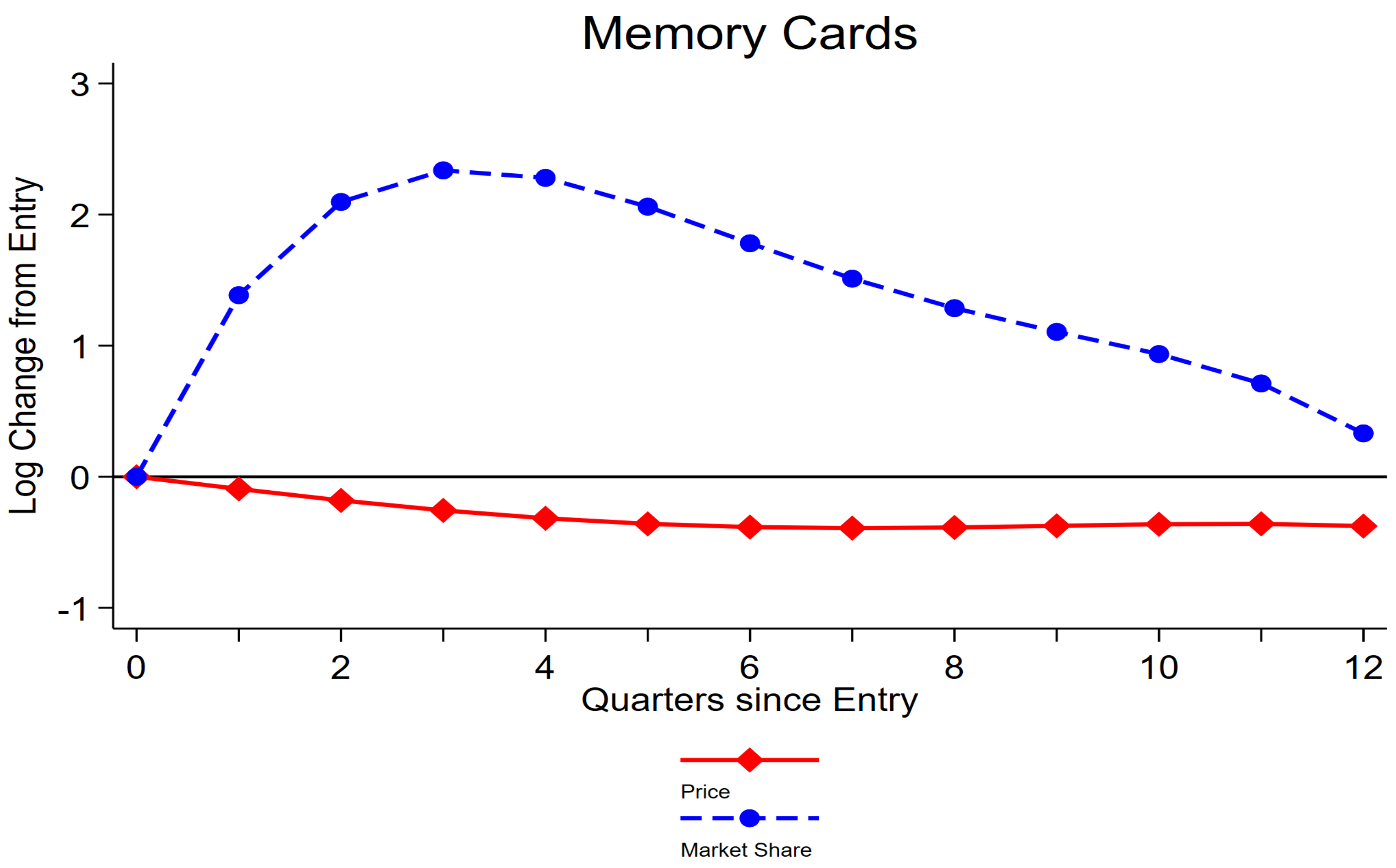
High turnover rates at item level: Some examples

Annualized Item-Level Entry and Exit Rates

	Entry Rate	Exit Rate
Memory Cards	25.3%	26.2%
Coffee Makers	24.8%	19.3%
Headphones	28.2%	23.9%
Boys' Jeans	54.6%	35.0%
Occupational Footwear	66.0%	49.6%

Source: "Quality Adjust at Scale" May 2023, Table 1

Inter-related Sales Share and Price Dynamics



Each point is relative to item-level value at entry.

Product group dominated by National Brands.

High pace of product entry and exit: 5.8% entry, 6.0% exit per qtr.

Rapid quality change: Substantial increases in speed and size.

Dynamic patterns vary substantially by product group

Current practice in the U.S. CPI

- Enumerators collect prices from retail outlets (in-person, phone, online)
- Conduct item substitution at stores when items are missing
  - Use “closest substitute” concept
  - Replacement prices can be compared, imputed, or quality adjusted
  - Requires attention of enumerator or analyst, sometimes in consultation with store managers
- Hedonic adjustments only applied to 7.5 percent of commodities
  - Requires coded attributes and good-by-good econometric modeling
- New items enter through outlet/item rotation on a <u>four-year cycle</u>

Current practice not feasible at scale

Challenges for implementing quality adjust in scanner data

- Millions of goods
- Rapid turnover mandates quality adjustment for all goods
- Need procedures that do not require analysis good-by-good
- Product characteristics might be not be encoded

Hedonic approaches

In general, predict price or price change as a function of attributes

- Entry and exit: Estimate virtual price for entering goods after they appear, or exiting goods after they disappear
- Continuing goods: Estimate change in value of attributes for unchanging goods

Machine Learning/AI essential for estimating value of millions of entering and exiting products

Assessing feasibility on inhaling unstructured data at scale: Grocery store example (NielsenIQ/Kilts data)

- Grocery, discount, convenience, drug and liquor store items for food and nonfood
- Weekly prices and quantities at store-item level, 2007-2015
- Hedonics:
  - Kilts data lacks coded attributes
  - Machine learning based on non-encoded, abbreviated product descriptions
  - Quality adjustment at scale for unstructured data
    - Training custom LLM for hedonic adjustment

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`

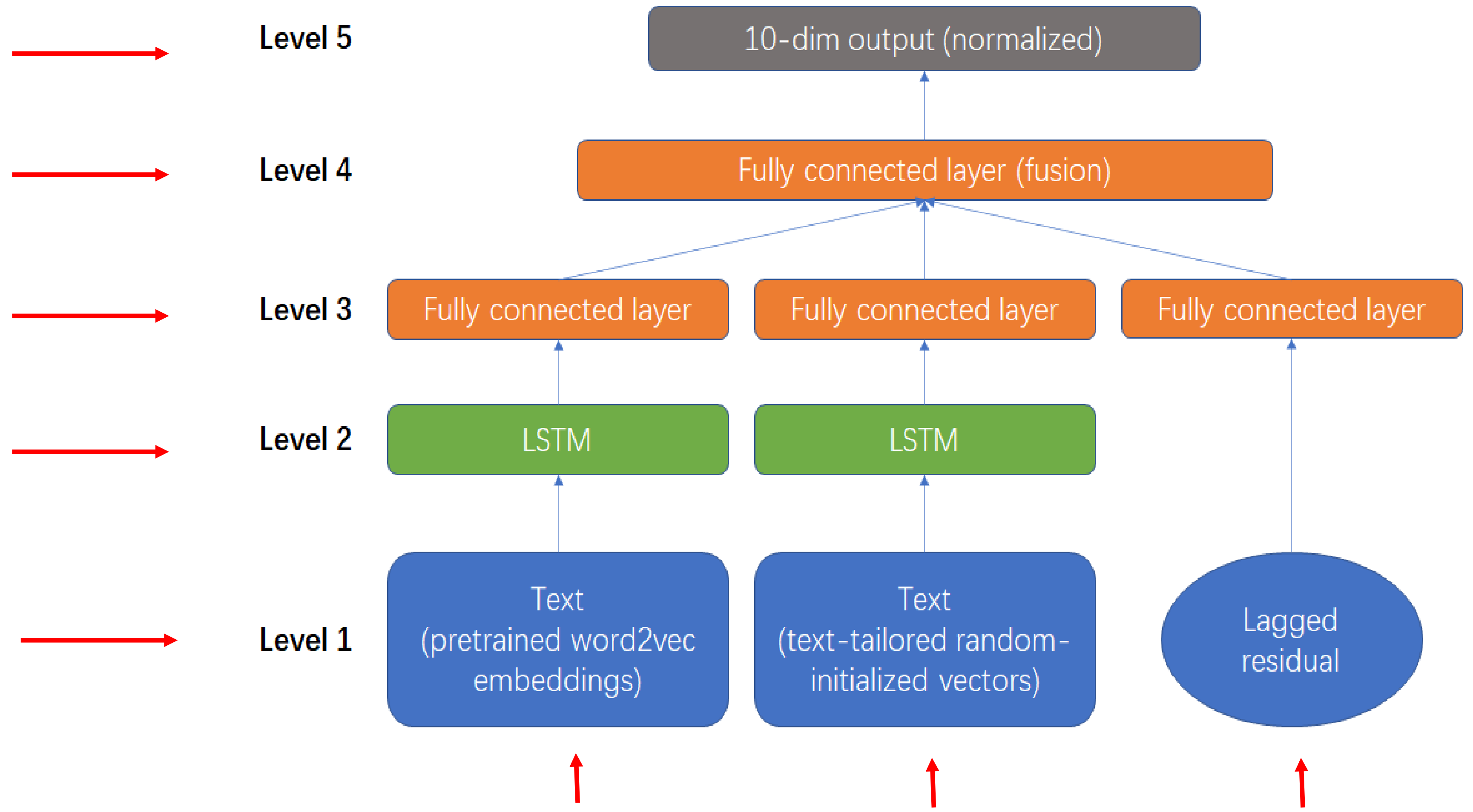
`1P`= 1 ply, `308S`= 308 sheets, `6PK`=6 pack

- Hiring a team to code would be very time and labor intensive.
- Instead we use machine learning…

Machine Learning Procedure

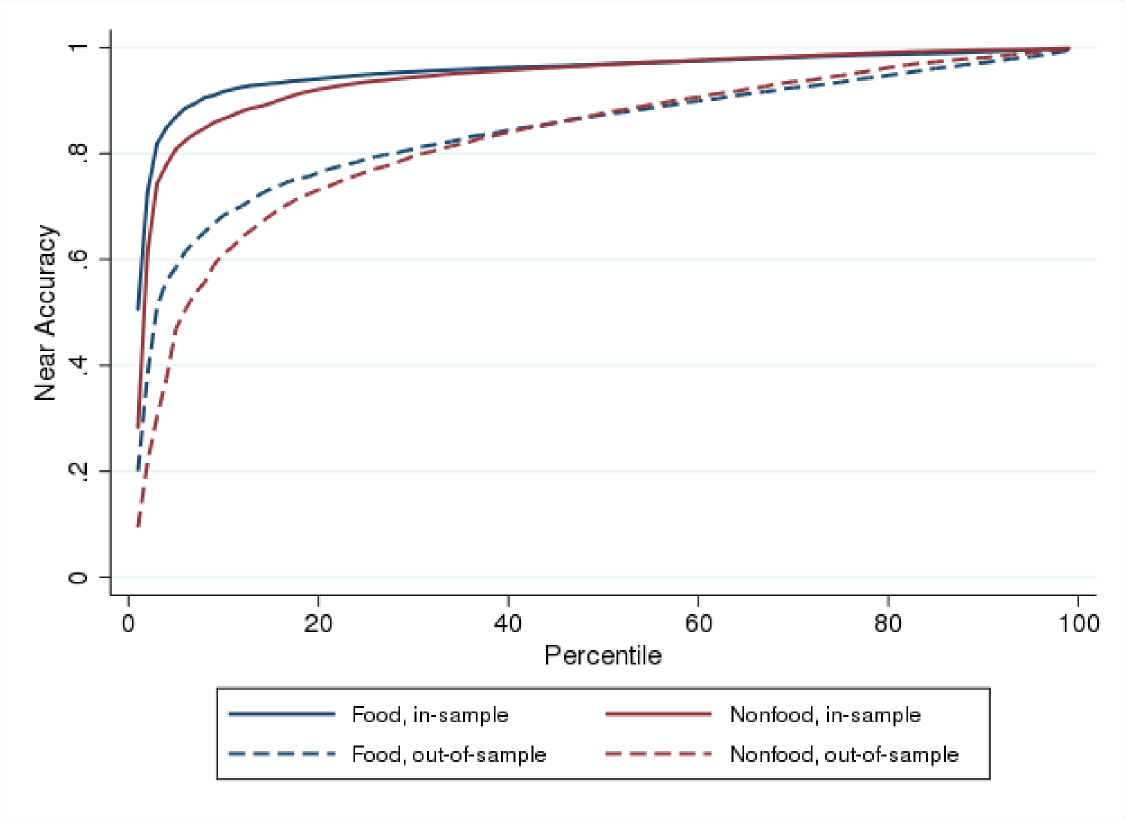
- Use natural language processing to create mappings between prices and characteristics for NielsenIQ using abbreviated text fields and brands.
- Integrate Erickson and Pakes approach to Hedonics into ML
  - First create predicted prices in log levels. Generate residuals
  - Repeat ML for price changes with lagged residual as additional embedding
- Cross-validation
  - Fit neural net model on 50% of data
  - Select model on 40% of data
  - Validate (report statistics) on 10% of data
- Binned prediction of prices and price changes (10 deciles)
- Fit continuous prices as in product of non-trivial bin probabilities with time/product-specific bin means

System Architecture (Recurrent Neural Network)

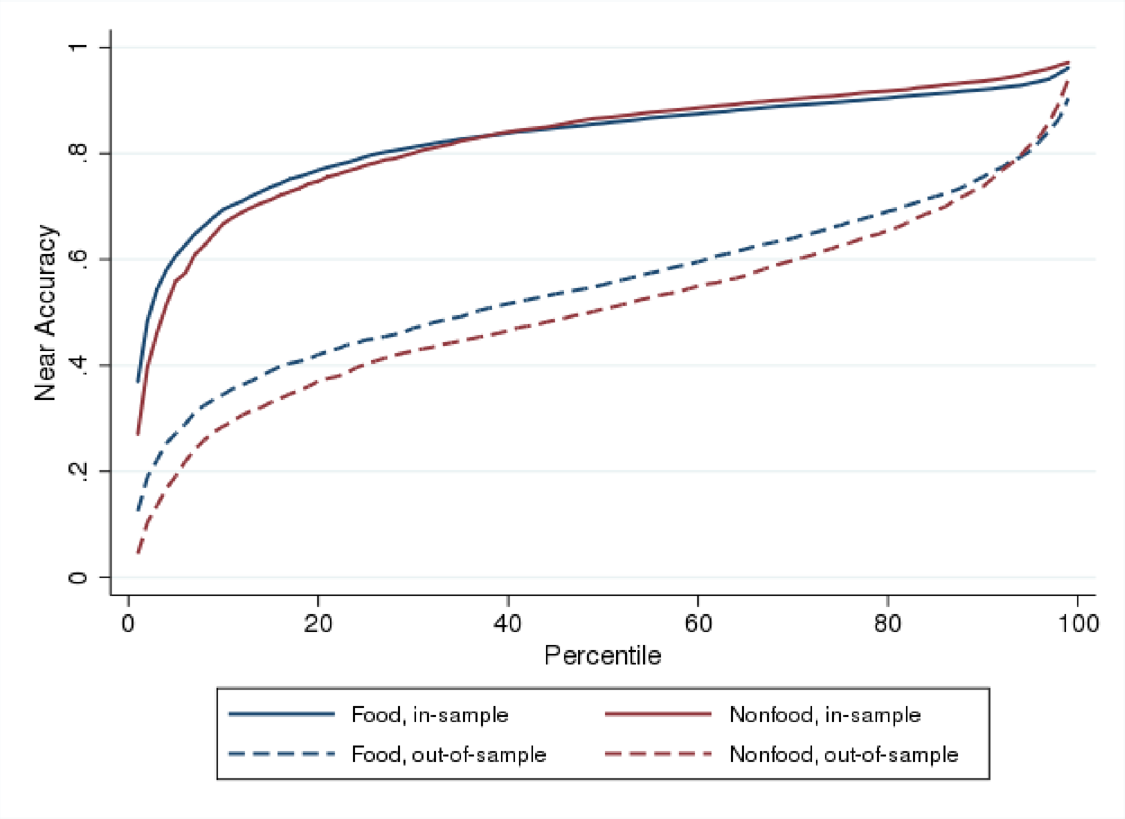


ML Model Fit: Near Accuracy

Correct or Adjacent Bin: Price Level



Correct or Adjacent Bin: Price Change



Model Fit: Near Accuracy, Selected food products

BABY FOOD	93.5%
FRESH PRODUCE	91.8%
COFFEE	90.4%
CARBONATED BEVERAGES	85.9%
BREAD AND BAKED GOODS	82.4%
PREPARED FOODS-FROZEN	80.3%
MILK	79.3%
CANDY	77.2%
SNACKS	77.2%
CEREAL	77.2%

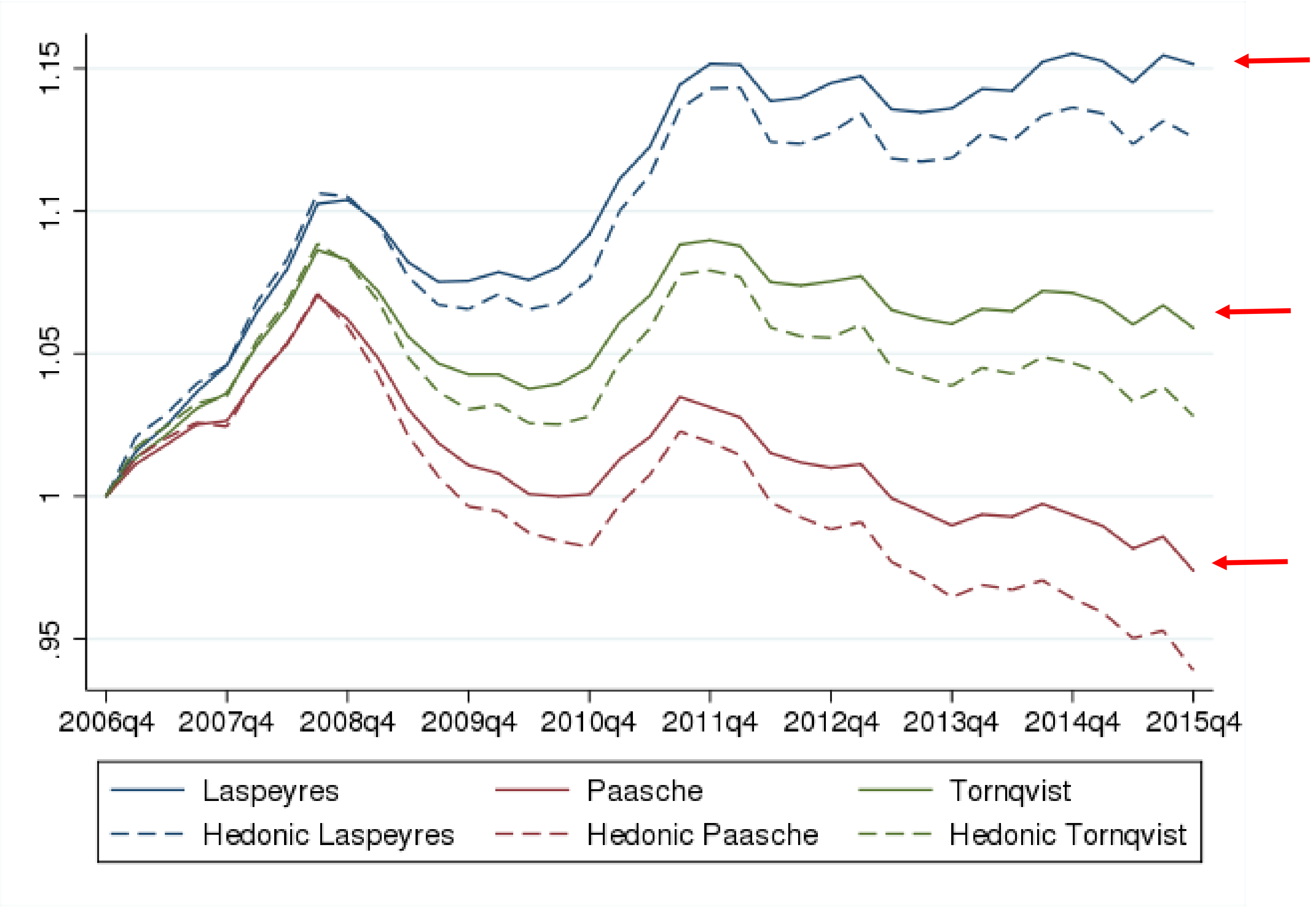
Price Indexes: Taking into account substitution effects

- Price indexes are weighted averages of price changes
- Laspeyres/Traditional CPI: Uses lagged weights
  - Fixed weigh neglects "substitution effect", Overstates inflation
- Paasche: Uses current weights
  - Overstates substitution; Understates inflation
- "Superlative indexes" (Tornqvist or Fisher): Averages weights across time; Gets substitution effects approximately correct
- **Requires simultaneous price and sales measurements**

Price Indexes:
Taking into account quality change with hedonics

- Using hedonics to predict prices from attributes:
  - Product exit:
    - What would product have sold for had it not disappeared?
  - Product entry:
    - What would product have sold for had it existed prior to entry?
- Superlative index formulas calculated over all goods using imputed price changes

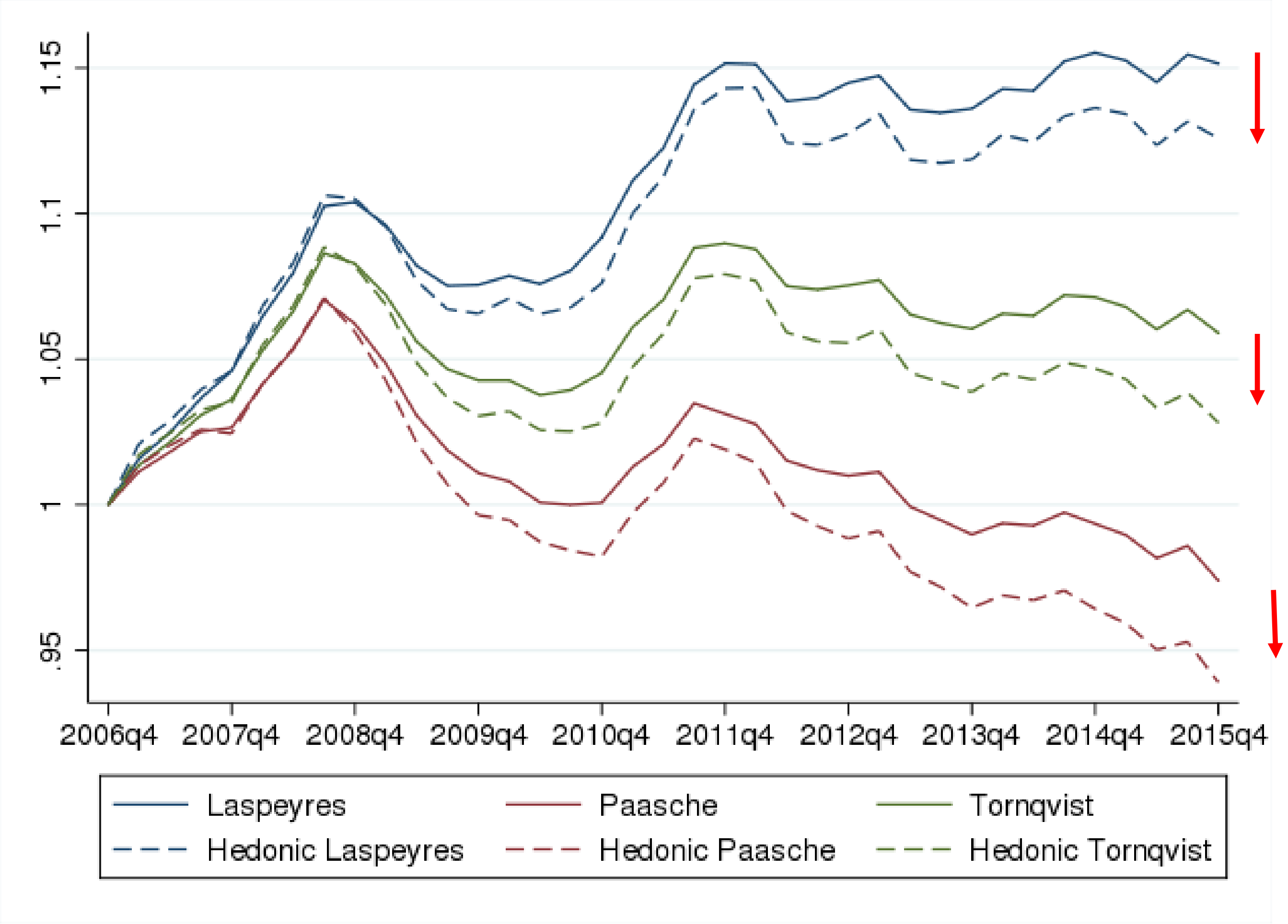
Price Indices – Food Product Groups



Substitution effects large in Food

- Traditional Tornqvist index indicates much lower inflation than Laspeyres or CPI: Substitution effects large
- Large substitution effects not surprising given big relative price movements in Food

Price Indices – Food Product Groups

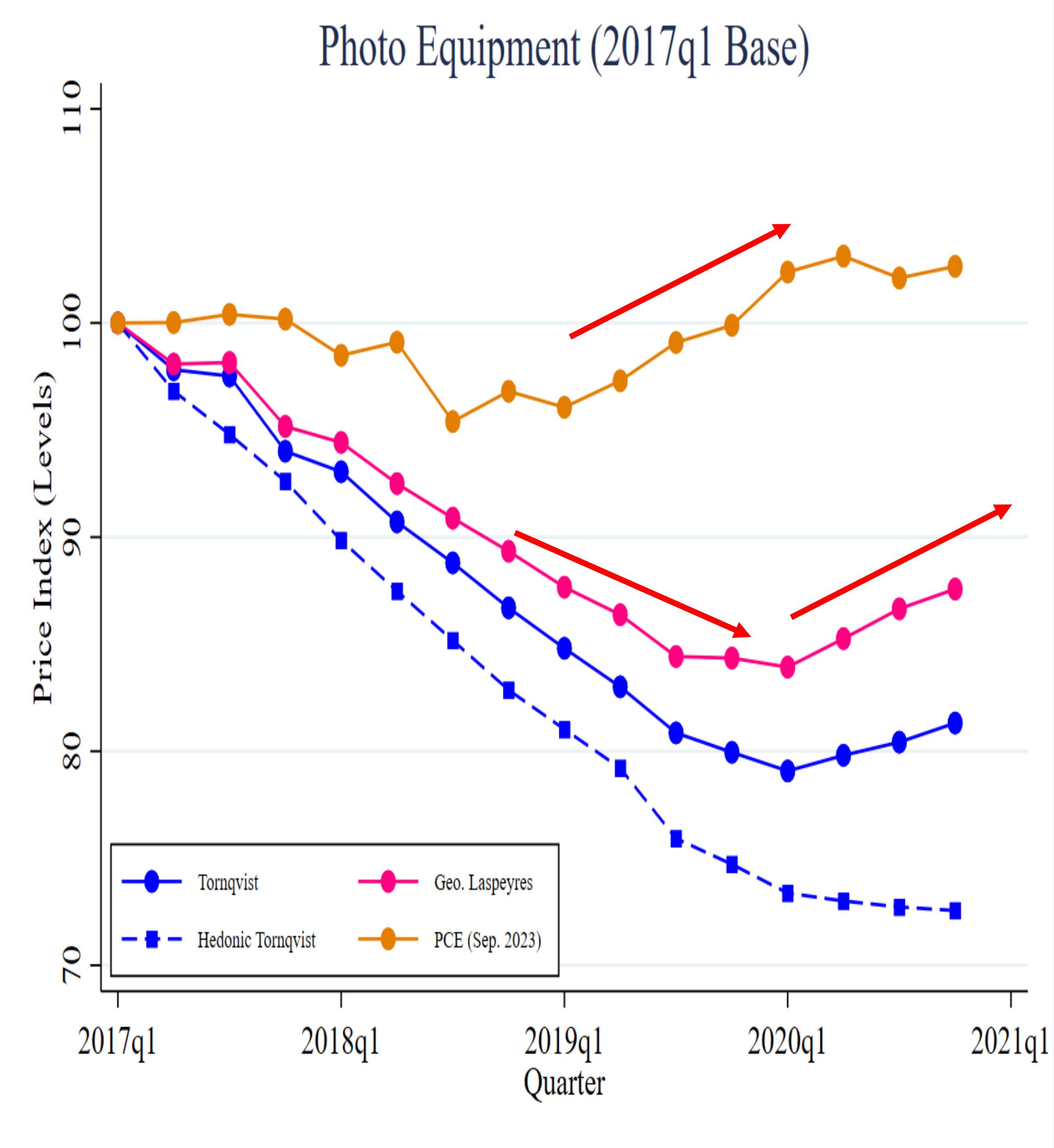


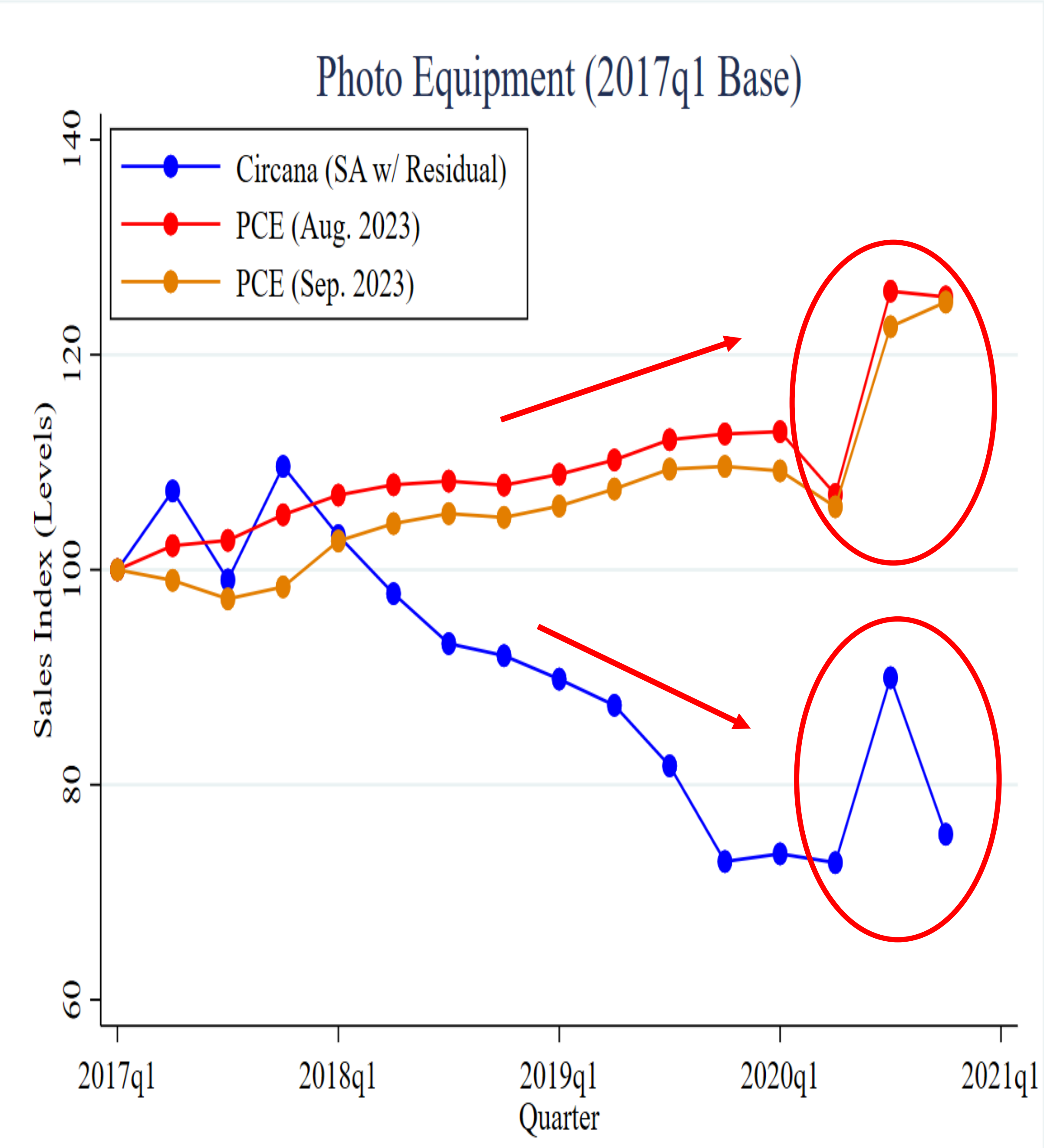
Substantial quality improvement—even in food

- Quality adjustment reduce food inflation by 0.25% per year
- ML techniques effective

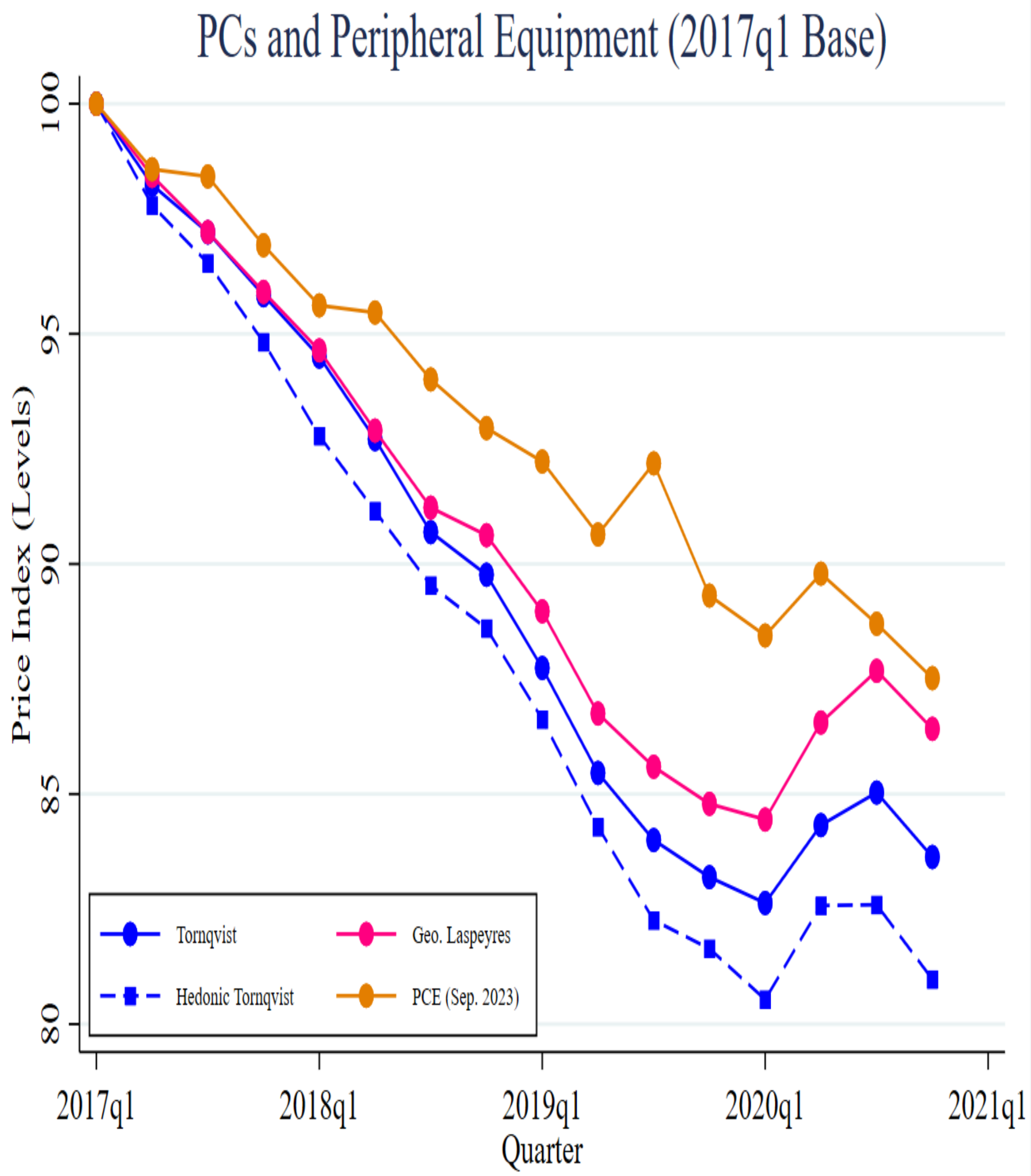
Nominal sales measures: Extrapolation/interpolation between Economic Censuses

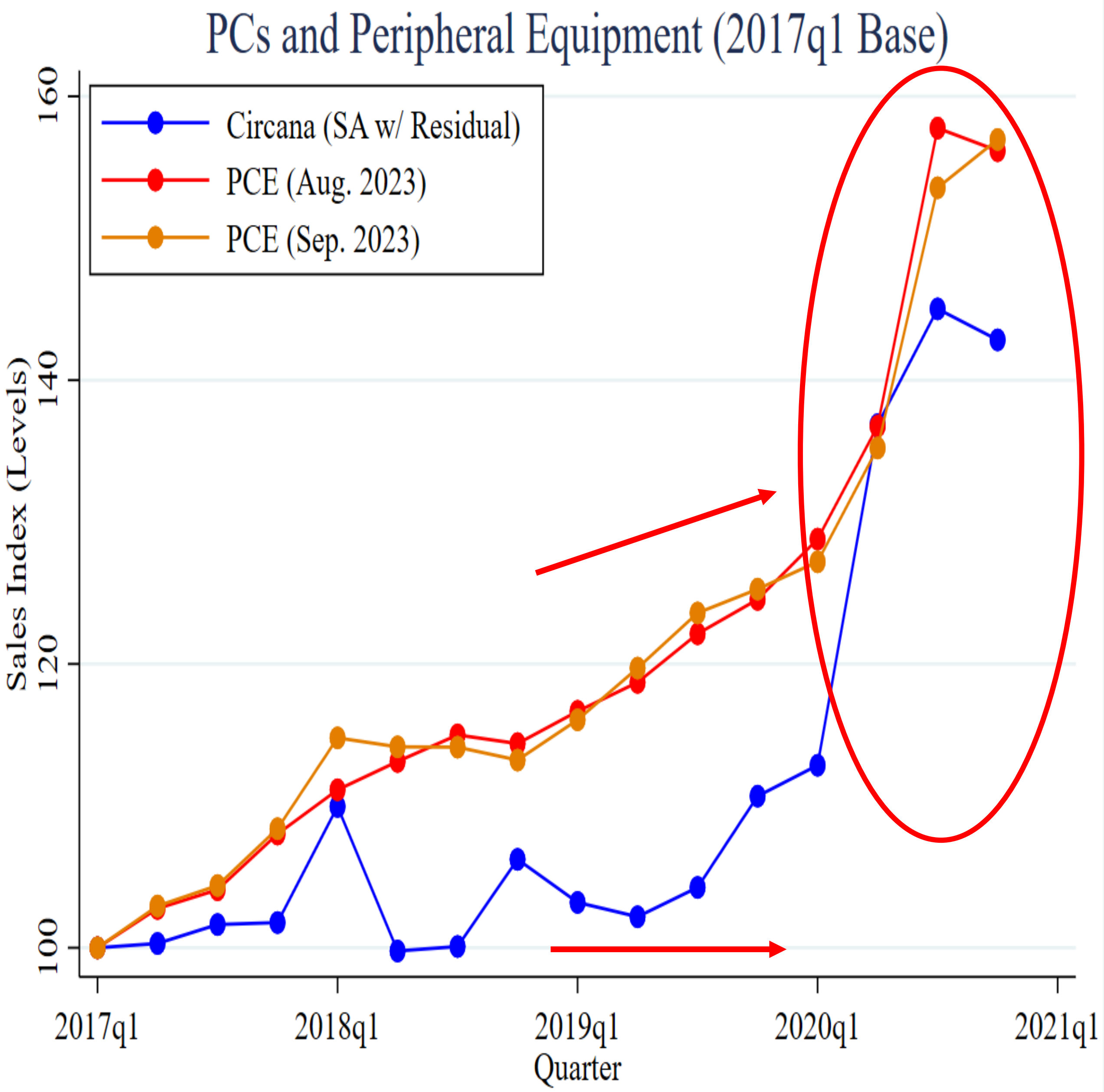
- Monthly Retail Sales
  - Backbone of high-frequency GDP
  - Measured at store/chain level
- Product detail collected in Economics Censuses (every 5 years)
- Monthly/Quarterly Personal Consumption Expenditures extrapolated and interpolations → huge errors in product detail





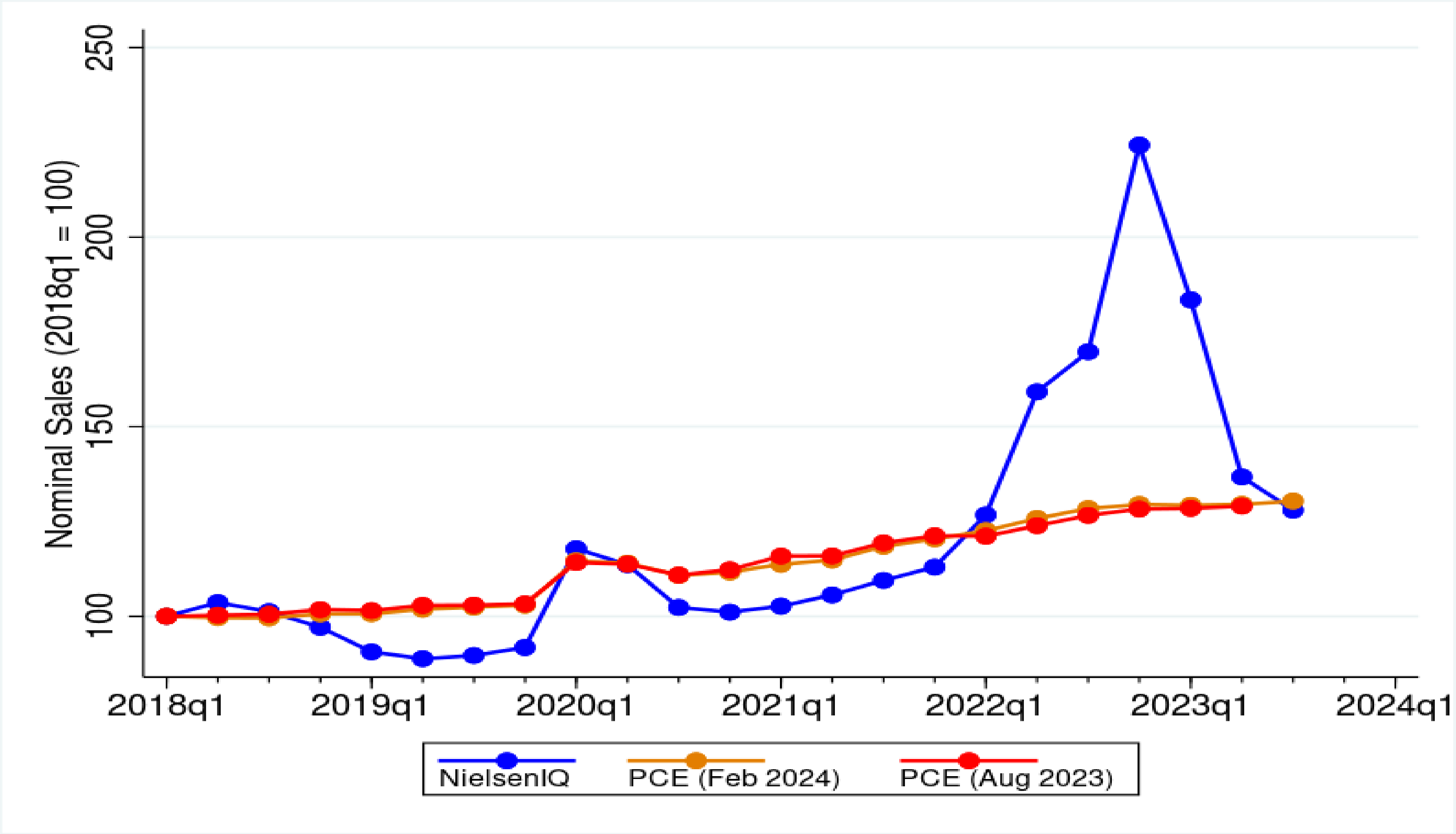
Sources: PCE data are from Bureau of Economic Analysis; other series authors calculations based on Circana data





Sources: PCE data are from Bureau of Economic Analysis; other series authors calculations based on Circana data

Nominal Sales of Eggs: Case of Bird Flu



Sources: PCE data are from Bureau of Economic Analysis; NielsenIQ data are authors' calculation

Lessons Learned

- Using item-level P and Q transactions data with attributes can be used to produce
  - Price indexes that adjust for substitution effects (superlative indexes)
  - Price indexes that adjust for product turnover and quality change
  - Quantitatively important:
    - Substitution effects and quality change roughly equally important
    - Under-measured quality change is ubiquitous, e.g., groceries
  - Necessary and feasible to do ubiquitous quality adjustment to deal with turnover in scanner data
- Implications
  - Inflation overstated on average
  - Productivity and real output growth understated
  - Compositional changes in inflation, sales mismeasured, e.g., COVID supply chain, recent inflation episode, extrapolation between Economic Censuses

RESET agenda for official statistics

- Expand methods to create new indicators at scale on timely basis
- Newly-funded RESET Demonstration Project
- Collaboration between RESET Project and Circana
  - Circana formed by merger of NPD and IRI
  - large aggregator of retail point-of-sales data (aka scanner data)
  - Item-level price and quantity: Store-level data to be aggregated nationally
  - Covers near-universe of transactions for “food at home,” “consumer packaged goods,” and “general merchandise” $\approx$ 2/3 of consumer goods
    - Includes food at home, health and beauty, apparel, electronics, tech good
    - Excludes automobiles, gasoline, and furniture

RESET Demonstration Project timeline and scope

- Coverage: near universe of transactions for majority of consumer goods (excluding vehicles, gasoline, furniture,and prescription drugs)
- Frequency: Monthly
- Timeliness: At least as timely as CPI and PCE
- Indicies: National-level indices of price and quantity
    - Consistent, simultaneous measures of price and quantity (instead of separate surveys from different sample)
    - Near census of transactions monthly (instead of 5-year Economic Census)

RESET Demonstration Project scope and timeline

Two types of data products

1. Series that use traditional formulas to mimic CPI and PCE
   - Are difference due to source data or approach?
   - Showing historically-comparable data necessary for acceptance/continuity
2. RESET measurement approach
   - Superlative indices consistently aggregated from item-level data
     - Improved estimates of effect of consumer substitution
     - Current approaches (based on high level of aggregation) understate substitution
   - Hedonics at scale
     - Do hedonic adjust of all goods, not just tech goods and apparel
     - Include all items (entering, continuing, and exiting); ubiquitous quality adjustment

RESET Demonstration Project scope and timeline

- Monthly releases
- Historical data beginning in 2020
- Look, feel, cadence of official statistical releases
- Time series published: National, Aggregate and product detail
- Notional timeline (ambitious)
    1. Food-at-home matched-model price indices: end of 2025
    2. Circana universe, matched-model price and quantity indices: 2026
    3. Hedonic adjustment, allowing inclusion of entry and exiting goods:
        - Quality adjustment as scle
        - Aim for 2027

Thank You & Additional Information

- The RESET Project's website is https://ebp-projects.isr.umich.edu/RESET/
- Papers:
  - Minding Your Ps and Qs: Going from Micro to Macro in Measuring Prices and Quantities
  - Re-engineering Key National Economic Indicators
  - Big Data for the 21st Century Economic Statistics: The Future is Now
  - Quality Adjustment at Scale: Hedonic versus Exact Demand-Based Price Indices
  - Using Machine Learning to Construct Hedonic Price Indices

Acknowledgements and Disclaimers

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

This presentation use researcher(s)' own analyses calculated (or derived) based in part on data from Nielsen Consumer LLC and marketing databases provided through the NielsenIQ Datasets at the Kilts Center for Marketing Data Center at The University of Chicago Booth School of Business. The conclusions drawn from the NielsenIQ data are those of the researcher(s) and do not reflect the views of NielsenIQ. NielsenIQ is not responsible for, had no role in, and was not involved in analyzing and preparing the results reported herein. Only University of Maryland and University of Michigan affiliates have worked with the NielsenIQ data via the Kilts Center.

This presentation also uses data from NPD/Circana housed at the U.S. Census Bureau. All results using the NPD/Circana data have been reviewed to ensure that no confidential information has been disclosed (CBDRB-FY19-122, CBDRB-FY21-074, CBDRB-FY23-067, CBDRB-FY23-0234). Opinions and conclusions expressed are those of the authors and do not necessarily represent the view of the U.S. Census Bureau.

We thank Neil Christy, Diyue Guo, Seula Kim, Jake Kramer, Michael Navarrete, Kelsey O'Flaherty, for their collaboration or assistance.

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 and Rebecca Hutchinson.

Technical slides

Price index formulas

Traditional Indices (geometric)

$$\ln \Phi^{L}_{t-1,t} = \sum_{k \in \mathbb{C}_{t-1,t}} s_{kt-1} \ln \frac{p_{kt}}{p_{kt-1}}$$

$$\ln \Phi^{P}_{t-1,t} = \sum_{k \in \mathbb{C}_{t-1,t}} s_{kt} \ln \frac{p_{kt}}{p_{kt-1}}$$

$$\ln \Phi^{TQ}_{t-1,t} = \sum_{k \in \mathbb{C}_{t-1,t}} \left(\frac{s_{kt-1} + s_{kt}}{2} \right) \ln \frac{p_{kt}}{p_{kt-1}}$$

- $\mathbb{C}_{t-1,t}$ is set of continuing goods from t-1 to t

Time-dummy method hedonics (used by BLS)

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

- Hedonic price index is δ_t
- Estimate separately for every pair of periods with Tornqvist weights (includes entering and exiting goods)
- Addresses entry/exit but not unobservable characteristics
- Imposes constant coefficients on characteristics in adjacent periods.

Erickson and Pakes (2011) Method

- Step 1: estimate hedonic function for log prices

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

- Step 2: estimate hedonic function for log price *changes* including lagged residual from first stage as an additional regressor:

$$\Delta \ln p_{kt} = g_t(Z_k, \hat{\eta}_{kt-1}) + \nu_{kt}$$

- Estimate both step 1 and step 2 relationships separately for each pair of periods
- Key insights:
  - Estimating price changes rather than levels differences out fixed unobserved characteristics (Pakes 2003)
  - Including the lagged residual helps to account for time-varying valuations of unobservable characteristics (Erickson and Pakes 2011)

Price index formulas: Generalized for entry/exit

Traditional Indices (geometric)

$$\ln \Phi^{L}_{t-1,t} = \sum_{k \in \mathbb{C}_{t-1,t}} s_{kt-1} \ln \frac{p_{kt}}{p_{kt-1}}$$

$$\ln \Phi^{P}_{t-1,t} = \sum_{k \in \mathbb{C}_{t-1,t}} s_{kt} \ln \frac{p_{kt}}{p_{kt-1}}$$

$$\ln \Phi^{TQ}_{t-1,t} = \sum_{k \in \mathbb{C}_{t-1,t}} \left(\frac{s_{kt-1} + s_{kt}}{2}\right) \ln \frac{p_{kt}}{p_{kt-1}}$$

Hedonic Indices (geometric)

$$\ln \Phi^{LH}_{t-1,t} = \sum_{k \in \mathbb{C}\mathbb{X}_{t-1,t}} s_{kt-1} \ln \frac{\widehat{p_{kt}}}{p_{kt-1}}$$

Exit

$$\ln \Phi^{PH}_{t-1,t} = \sum_{k \in \mathbb{C}\mathbb{E}_{t-1,t}} s_{kt} \ln \frac{\widehat{p_{kt}}}{p_{kt-1}}$$

Entry

$$\ln \Phi^{TQH}_{t-1,t} = \sum_{k \in \mathbb{C}\mathbb{E}\mathbb{X}_{t-1,t}} \left(\frac{s_{kt-1} + s_{kt}}{2}\right) \ln \frac{\widehat{p_{kt}}}{p_{kt-1}}$$

Exit & Entry

- $\mathbb{C}_{t-1,t}$ is set of continuing goods from t-1 to t; $\mathbb{X}_{t-1,t}$ is exiting goods, $\mathbb{E}_{t-1,t}$ is entering goods,
- Hedonic Laspeyres adjusts for exit; Hedonic Paasche adjusts for entry
- Hedonic Tornqvist adjusts for entry and exit