



Using Naturally-Occurring Data for  
Retail Sales, CPI, and PCE:  
The Future is Now

Matthew D. Shapiro  
University of Michigan and NBER

Presentation at the  
Federal Economic Statistics Advisory Committee  
December 9, 2016

# Naturally-occurring or non-designed data for consumer spending and prices

Household transactional data, e.g.,

- Michigan-Berkeley account data project
- JPMorgan Institute
- Homescan

# Naturally-occurring or non-designed data for consumer spending and prices

## Aggregated transaction data

- Credit/debit card transactions
  - FRB/Palantir
  - BEA pilot
- Non-retail transactions
  - Hotel
  - Airlines
  - Movie Theater
  - Medical

# Naturally-occurring or non-designed data for consumer spending and prices

Price data, e.g.,

- Scanner data, e.g., Nielsen
- CPI pilot (presentation today)
- Webscraped, e.g., Billion Prices Project
- Redding-Weinstein project (last meeting)

# Naturally-occurring or non-designed data for consumer spending and prices

Sales and unit value data combined

- *Scanner data, e.g., Nielsen\**
- *CPI pilot (presentation today)\**
- Retailer transactions

\*Joint price, sales measurement not implemented

# Retail transactions data

- Detailed, SKU level
  - Sales
  - Unit values
- Aggregated to ELI-like level
  - Sales
  - Price indexes
  - Joint measurement of price and quantity
- Transmitted to statistical agencies
  - FRB/Palantir software tool

# Current Architecture

<p><b>Census (nominal spending)</b></p> <p><b>Data collection:</b> Retail Trade surveys (monthly and annual) Economic Census (quinquennial)</p> <p><b>Published statistics:</b> Retail Trade (monthly)</p>	<p><b>BLS (prices)</b></p> <p><b>Data collection:</b> Consumer expenditure survey (spending weights) Telephone Point of Purchase survey (purchase location) CPI price enumeration (Probability sampling of goods within outlets)</p> <p><b>Published statistics:</b> Consumer Price Index (monthly)</p>
<p style="text-align: center;"><b>BEA (aggregation and deflation)</b></p> <p><b>Data collection:</b> Census and BLS data supplemented by multiple source</p> <p><b>Published statistics:</b> Personal Consumption Expenditure: Nominal, real, and price (monthly) GDP (quarterly)</p>	

# Current Architecture

<p><b>Census (nominal spending)</b></p> <p>Retail Trade surveys Economic Census</p> <p>→nominal sales</p>	<p><b>BLS (prices)</b></p> <p>Consumer expenditure survey →weights Telephone Point of Purchase survey →outlets CPI price enumeration →price quotations</p> <p>→price indexes</p>
<p><b>BEA (aggregation and deflation)</b></p> <p>Personal Consumption Expenditure: Nominal, real, and price</p>	



# New data: Household Accounts

<p><b>Census (nominal spending)</b></p> <p>Retail Trade surveys Economic Census</p> <p>→nominal sales</p>	<p><b>BLS (prices)</b></p> <p>Consumer expenditure survey →weights Telephone Point of Purchase survey →outlets CPI price enumeration →price quotations</p> <p>→price indexes</p>
<p><b>BEA (aggregation and deflation)</b></p> <p>Personal Consumption Expenditure: Nominal, real, and price</p>	

# New data: Household Accounts

## Census (nominal spending)

Retail Trade surveys  
Economic Census

→ **nominal sales**

## BLS (prices)

Consumer expenditure survey → **weights**

Telephone Point of Purchase survey → **outlets**

CPI price enumeration → price quotations

→ price indexes

## BEA (aggregation and deflation)

Personal Consumption Expenditure: **Nominal**, real, and price

# New data: Transaction aggregators

<p><b>Census (nominal spending)</b></p> <p>Retail Trade surveys Economic Census</p> <p>→nominal sales</p>	<p><b>BLS (prices)</b></p> <p>Consumer expenditure survey →weights Telephone Point of Purchase survey →outlets CPI price enumeration →price quotations</p> <p>→price indexes</p>
<p><b>BEA (aggregation and deflation)</b></p> <p>Personal Consumption Expenditure: Nominal, real, and price</p>	

# New data: Transaction Aggregators

<p><b>Census (nominal spending)</b></p> <p>Retail Trade surveys Economic Census</p> <p>→ <b>nominal sales</b></p>	<p><b>BLS (prices)</b></p> <p>Consumer expenditure survey → <b>weights</b></p> <p>Telephone Point of Purchase survey → <b>outlets</b></p> <p>CPI price enumeration → price quotations</p> <p>→ price indexes</p>
<p><b>BEA (aggregation and deflation)</b></p> <p>Personal Consumption Expenditure: <b>Nominal</b>, real, and price</p>	

# New data: Web scraped prices

<p><b>Census (nominal spending)</b></p> <p>Retail Trade surveys Economic Census</p> <p>→nominal sales</p>	<p><b>BLS (prices)</b></p> <p>Consumer expenditure survey →weights Telephone Point of Purchase survey →outlets CPI price enumeration →price quotations</p> <p>→price indexes</p>
<p><b>BEA (aggregation and deflation)</b></p> <p>Personal Consumption Expenditure: Nominal, real, and price</p>	

# New data: **Web scraped prices**

<p><b>Census (nominal spending)</b></p> <p>Retail Trade surveys Economic Census</p> <p>→nominal sales</p>	<p><b>BLS (prices)</b></p> <p>Consumer expenditure survey →weights Telephone Point of Purchase survey →outlets CPI price enumeration →<b>price quotations</b> →<b>price indexes</b> (with external weights)</p>
<p><b>BEA (aggregation and deflation)</b></p> <p>Personal Consumption Expenditure: Nominal, real, and price</p>	

# New data: Retail transactions

<p><b>Census (nominal spending)</b></p> <p>Retail Trade surveys Economic Census</p> <p>→ nominal sales</p>	<p><b>BLS (prices)</b></p> <p>Consumer expenditure survey → weights Telephone Point of Purchase survey → outlets CPI price enumeration → price quotations</p> <p>→ price indexes</p>
<p><b>BEA (aggregation and deflation)</b></p> <p>Personal Consumption Expenditure: Nominal, real, and price</p>	

# New data: Retail transactions

<p><b>Census (nominal spending)</b></p> <p>Retail Trade surveys Economic Census</p> <p>→ nominal sales</p>	<p><b>BLS (prices)</b></p> <p>Consumer expenditure survey → weights</p> <p>Telephone Point of Purchase survey → outlets</p> <p>CPI price enumeration → price quotations (unit values)</p> <p>→ price indexes</p>
<p><b>BEA (aggregation and deflation)</b></p> <p>Personal Consumption Expenditure: Nominal, real, and price</p>	



# New data: Retail transactions

<p><b>Census (nominal spending)</b></p> <p>Retail Trade surveys Economic Census</p> <p>→ <b>nominal sales</b></p>	<p><b>BLS (prices)</b></p> <p>Consumer expenditure survey → <b>weights</b></p> <p>Telephone Point of Purchase survey → <b>outlets</b></p> <p>CPI price enumeration → <b>price quotations</b> (unit values)</p> <p>→ <b>price indexes</b></p>
<p><b>BEA (aggregation and deflation)</b></p> <p>Personal Consumption Expenditure: <b>Nominal, real, and price</b></p>	

# New Architecture: Retail transactions

- Integrates price and quantity measurement
- Combines multiple data collections
  - Retail sales survey
  - CPI: Multiple data collections
- Potential measurement improvements
  - Timeliness
  - Frequency
  - Geographical detail
  - Accounting for changing goods

# New Architecture: Challenges

- Requires retailer cooperation
- Turnover of goods and services
- New techniques for constructing price indexes
  - Revealed preference approach (Feenstra/Redding-Weinstein)
  - New hedonics, aided by machine learning
- Challenges for statistical agencies
  - Technical
  - Organizational