Using "Big Data" to Study Consumer Spending in Real Time

Claudia Sahm
Federal Reserve Board

May 18, 2018

The views expressed here are those of the author and not necessarily those of other members of the Federal Reserve System.

What's the Big Deal with "Big Data"?

Can innovations in big data – more speed and detail—improve macro policy making?
summary of Q&A remarks from Fed Chair Janet Yellen at Stanford, starts minute 39

Promise

- Huge amount of detailed firm-level data created in business operations
- Address gaps in official statistics, geographic or high frequency

Challenge

- Not constructed with statistical rigor and theory as official statistics
- Short time series, not representative, often not comparable to prior events

Interesting Examples to Study on Macro Events

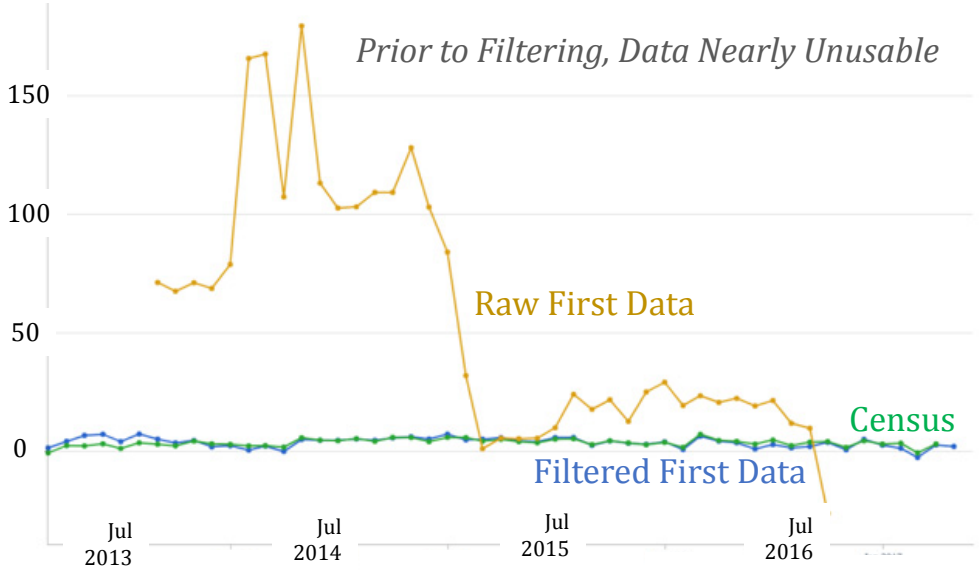
- "The Effect of Hurricane Matthew on Consumer Spending"
  Aladangady, Aron-Dine, Dunn, Feiveson, Lengermann, and Sahm
- "The Response of Consumer Spending to Changes in Gasoline Prices"
  Gelman, Gorodnichenko, Kariv, Koustas, Shapiro, Silverman, and Tadelis

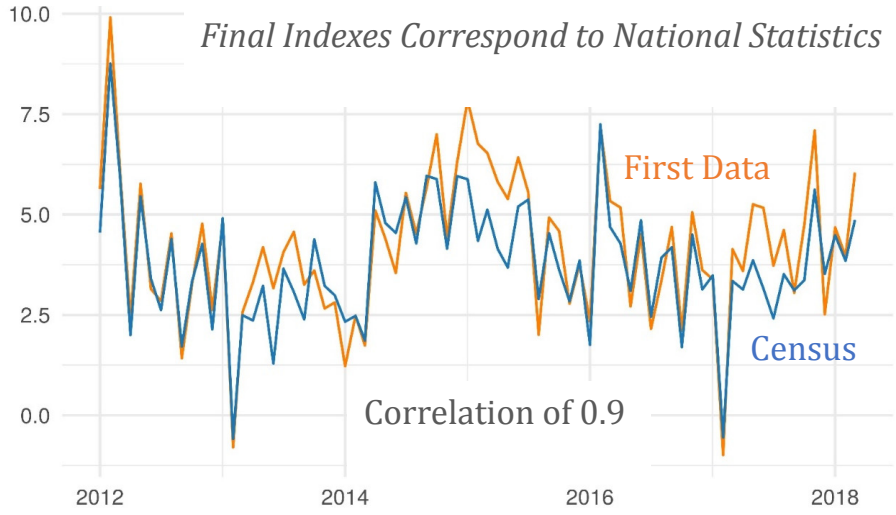
Collaboration: "Big Data" is a Big TEAM effort

- Finding ways for private companies to share data created in doing business for economic statistics and research is an important public good
- Collaborating with First Data, a large, payments processor with more than $2 trillion in card transactions per year across 6 million merchant locations
- High-quality Big Data is a multi-disciplinary effort: engineers and programmers at Palantir combine and filter merchant-level transactions, output anonymized indexes; economists at the Board develop filters, validate, use anonymized indexes in analysis
- Comparisons with existing data require working with statistical agencies like Census Bureau and Bureau of Economic Analysis

Filtering and Validation Prior to Analysis

12-month percent change, spending at retail stores and restaurants





FILTER

- Focus on economic changes in spending and exclude changes specific to business operations
- Palantir and Board defined "14-month chained merchant" filter for anonymized, time series

VALIDATE

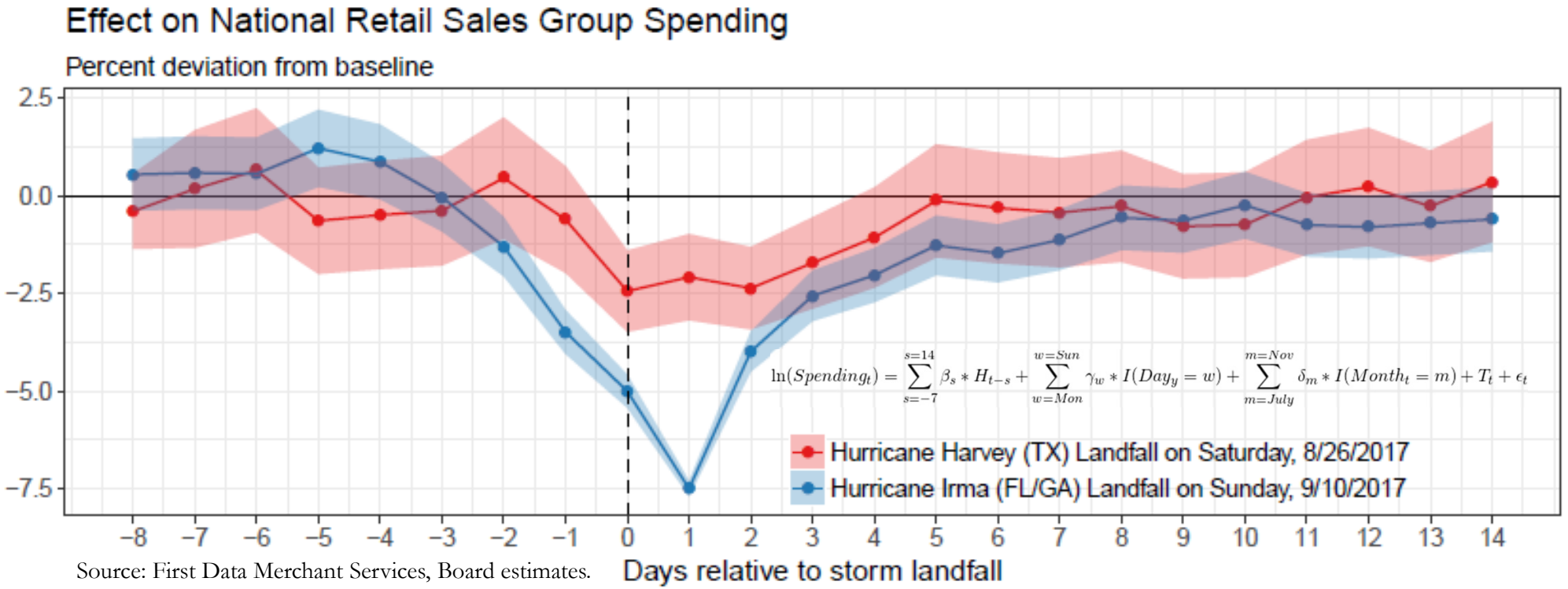
- Data set constructed with industry structure (4-digit NAICs) as Census Retail Trade Survey
- Compare sales at retail stores and restaurants (durables ex autos, nondurables ex gas, plus food services, about 25% of GDP) to national, monthly Census retail sales data

Study of Hurricanes Harvey and Irma in Real Time

Source: First Data Merchant Services, National Oceanic and Atmospheric Administration (NOAA); Board estimates.

- Spending from First Data available within **only three days**
- Using methods from our earlier study, tracked and reported in real time

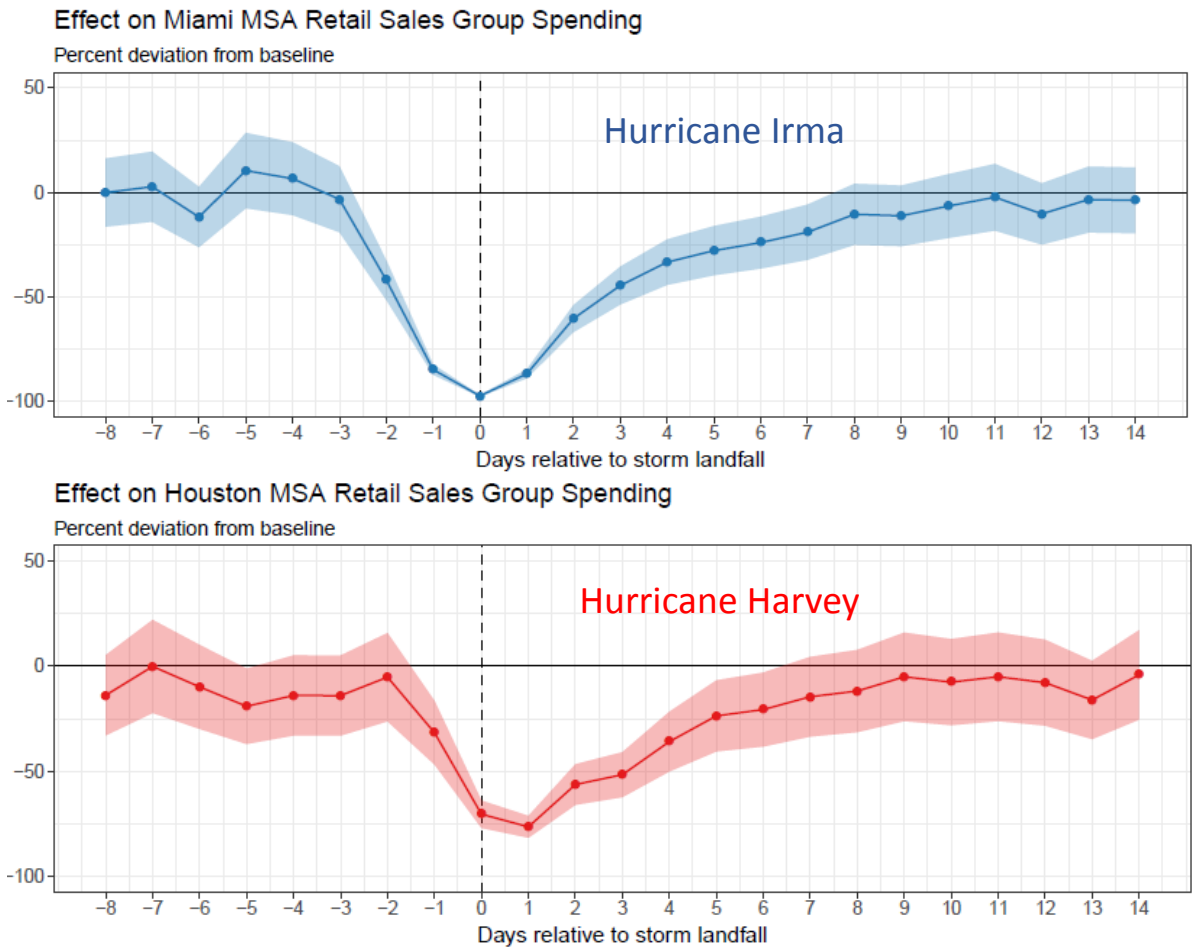
Decline in National, Daily Spending with Hurricanes



Source: First Data Merchant Services, Board estimates.

- Effect on national spending depends on population in storm's path
- Within week, spending back to normal but no rapid make up in spending

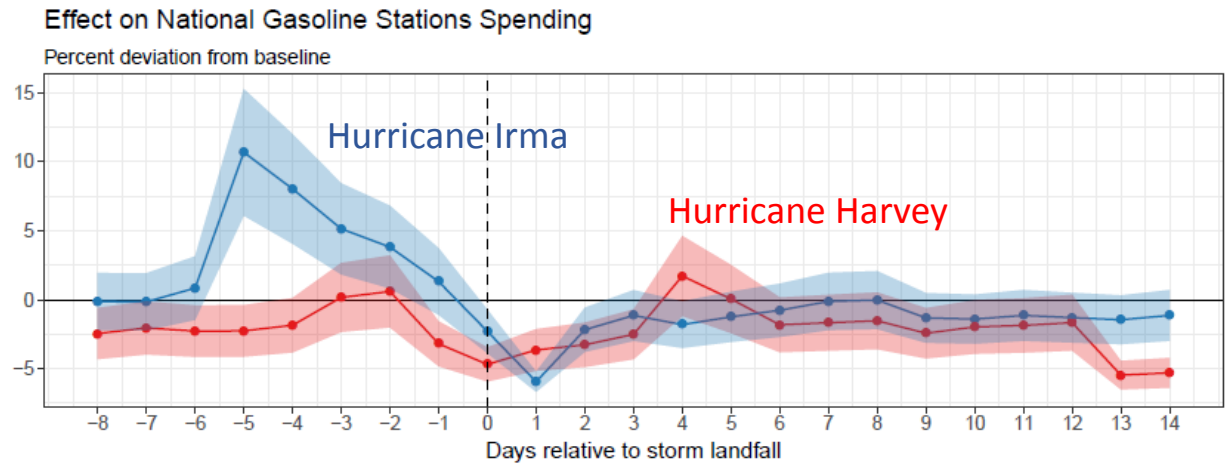
Spending in Direct Path of Hurricanes Nearly Halted

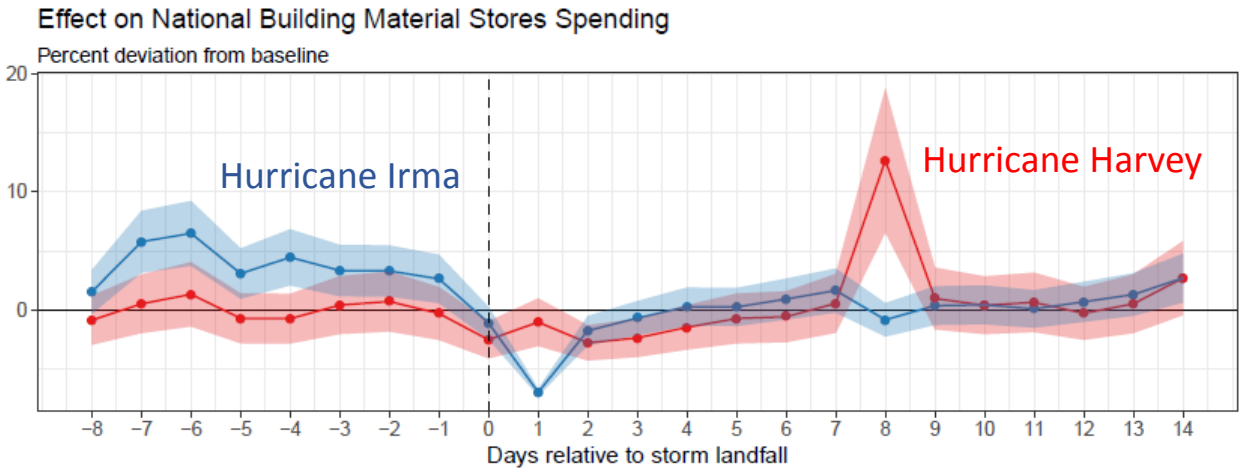


- Almost 100% decline in Miami (and also Tampa), may reflect power outages; 75% decline in Houston
- Before storm varies, may relate to advance warning or evacuation orders
- No sign of rapid makeup in disrupted spending

Source: First Data Merchant Services, Board estimates.

Spending on Gas and Building Materials Also Affected





- Large jump in gas spend before Irma coincides with evacuation orders
- Building materials spiked in percent on Sunday (lowest volume day of the week) after Harvey flood waters receded

Source: First Data Merchant Services, Board estimates.

Translating Daily Spend to Quarterly GDP

- **Direct estimate**: sum percentage "deviation from baseline" in daily retail group spending both hurricanes (slide 6), divide by 92 days in quarter, scale by spend category's share of GDP (about 0.25)

  implies **almost ½ percentage point less GDP growth** in 17Q3

- Gradual makeup (unlike sharp drop) hard to distinguish from usual variability, so direct estimate **may overstate negative effect**

- Estimate from **retail stores excludes** other consumption, like recreation services, or unplanned inventory accumulation or production disruptions, see also Bayard, Decker, and Gilbert (2017)

- **May be hard to see effect in official statistics** on retail sales. National sampling frame may not measure localized shocks well

Concluding Remarks

- Only **at early stages** of harnessing "Big Data" for macro policy analysis; requires new techniques and carefully unpacking new variation
- "Big Data" shed light on **high-frequency, localized shocks in real time**, such as hurricanes, and may help fill gaps in official national statistics
- New variation in "Big Data" may help distinguish between competing **models of consumer behavior;** improve forecasts from other shocks
- **No single type of data can do it all**. We need "Big Data," administrative data, as well many surveys of firms and households