

Big Data and Security

Jeffrey Borowitz, PhD

Lecturer

Sam Nunn School of International Affairs

Crawled Price Data

Price Data

- Data on prices in particular is readily available online
 - Companies post prices of goods in order to sell them
- One important economic indicator is inflation, which is a measure of the change in price indicators over time



Source: U.S. Department of Labor: Bureau of Labor Statistics

2014 research.stlouisfed.org

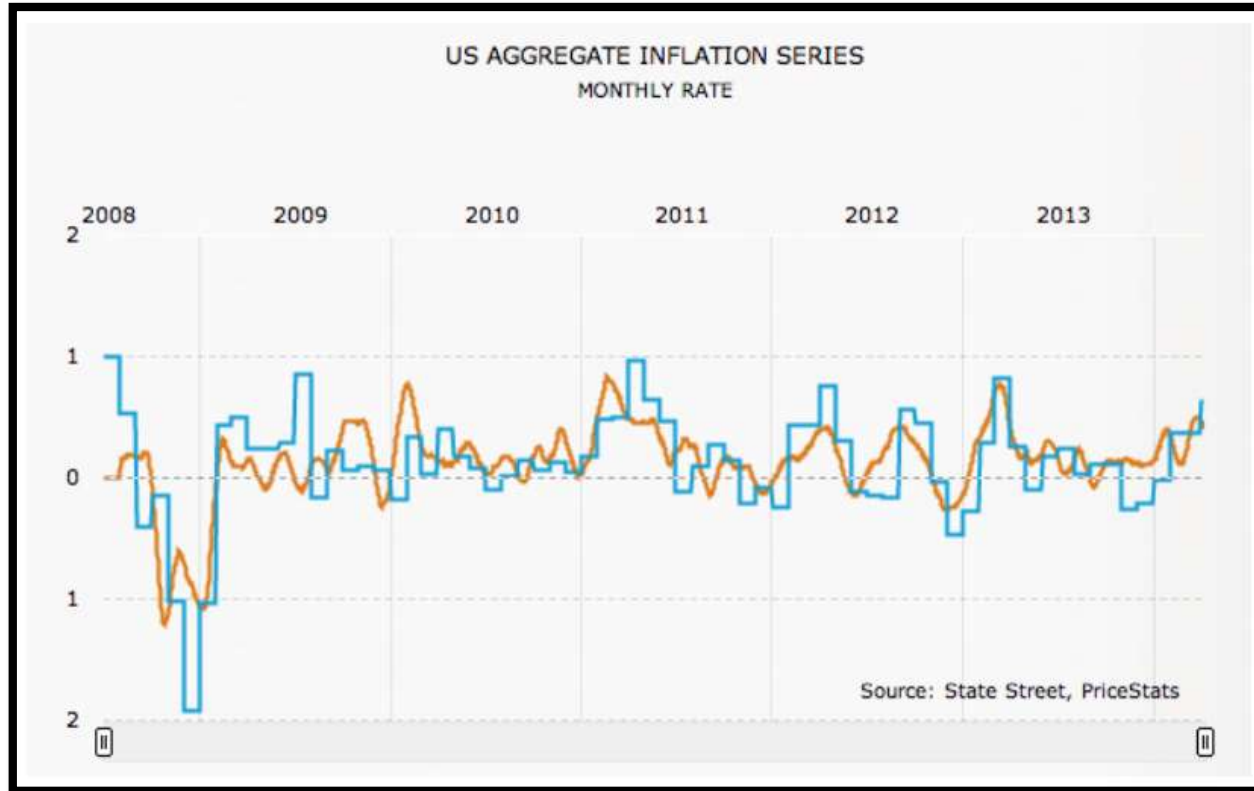
Inflation Methodology

- The Bureau of Labor Statistics
 - Collect prices for particular goods in particular places
 - E.g. electricity in Boston Use surveys of companies at “point of sale”
- Average across cities and across time
 - But how do you weight e.g. electricity and milk in Boston?
 - Weight equally?
 - Weight based on what people pay (Consumer Price Index, based on the Consumer Expenditure Survey)
 - Based on what companies pay (Producer Price Index, comes from BEA data)

Big Data and Inflation

- Big data makes it easy to collect price information, for a subset of goods
 - Online retail is particularly easy!
 - The Billion Prices Project (academic research at MIT), now a company called PriceStats.

PriceStats Price Index



PriceStats ©. (n.d.). Retrieved from <https://www.pricesstats.com/inflation-series?chart=1837>

PriceStats Coverage

Missing are places with relatively few people/relatively poor people



PriceStats ©. (n.d.). Retrieved from <https://www.pricestats.com/approach/data-composition>

Could PriceStats replace CPI?

- PriceStats only does something different to collect prices for goods in particular places
- They don't actually solve the weighting problem
- So big data can't completely replace traditional methodology this way
- But it could replace the point of sale survey
- Though it might not work as well for producer prices, etc

Premise

- Premise is a similar startup to PriceStats, but also uses local data from employees who take photos of stores and markets
- <http://www.premise.com/>
- Compared to PriceStats
 - They can conceivably measure “quality” a little better
 - They can measure things offline
 - And they can see when things are out of stock (empty shelves)

Prospects for Big Data and CPI

- The big hole in producing CPI using only big data is missing the weights.
- But where could we find estimates of the weights?
- We need info on what people are buying!
- Two options:
 - Big retailers
 - Credit card networks/payment processors
- These have their issues too, but are a step in the right direction

Epilogue

- Premise has become a “ground truth platform”
 - They use their crowdsourced people/network to sell specific information that companies need collected
- PriceStats struck a deal with State Street to sell their series
 - They also created a purchasing power parity index

Lesson Summary

- Inflation is computed by surveys of spending and prices
- Measures of prices can be updated with spending