Statistical Office and Big Data
“Blending into the future”

Roberto Rigobon

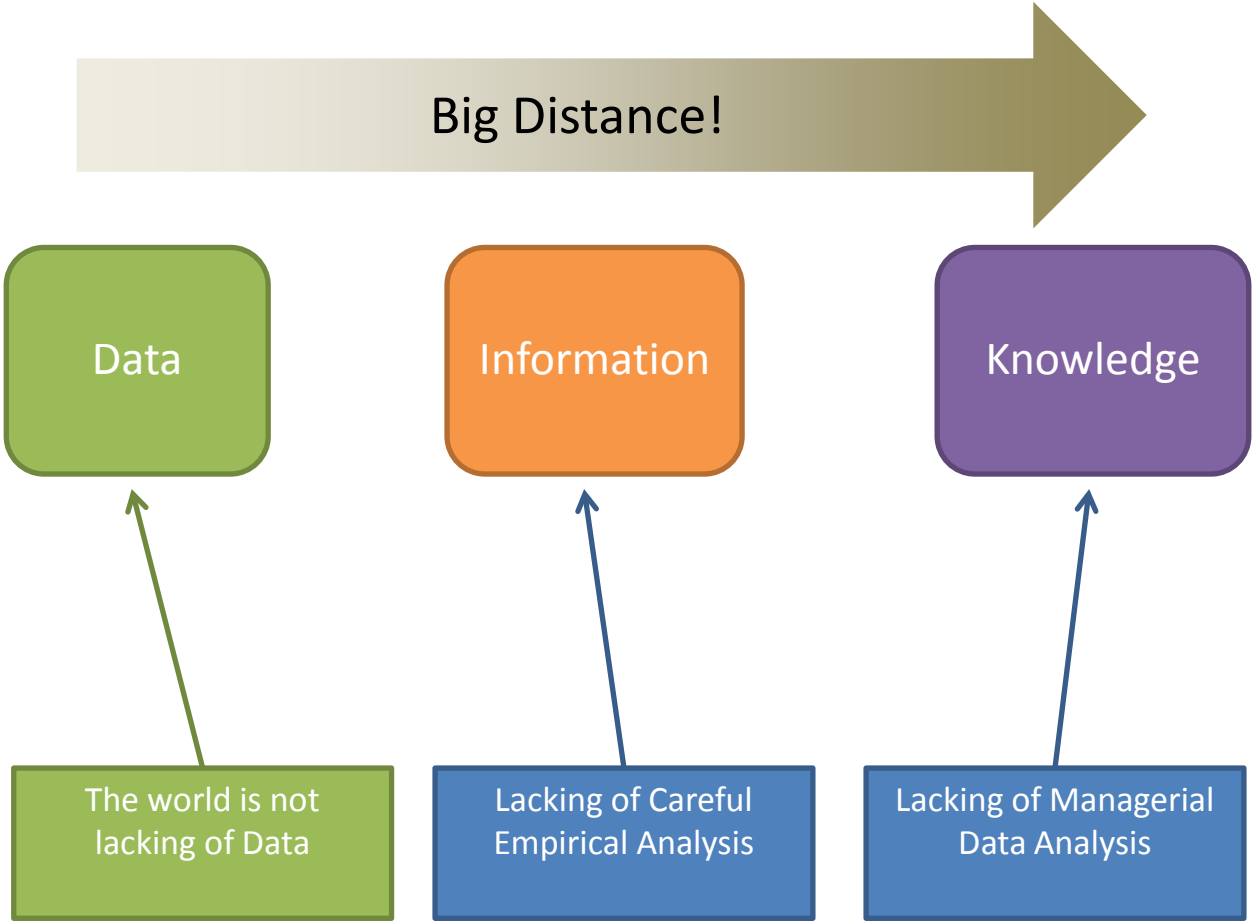
MIT, CSAC, CNStat, NBER

MIT MANAGEMENT SLOAN SCHOOL

Big Data

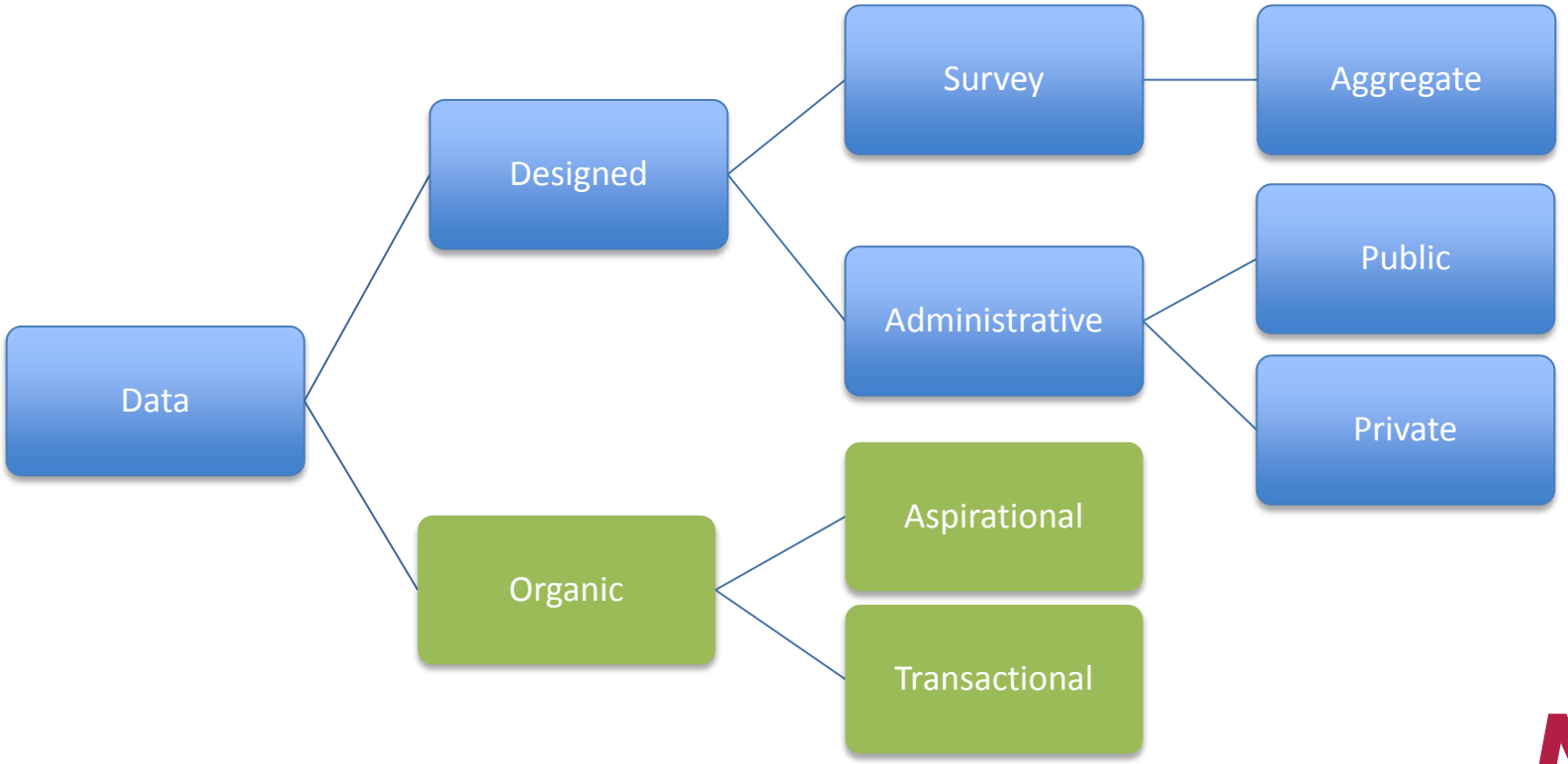
- Opportunity
  - New technologies and consumer practices allows us to expand the data collected
    - Timely
    - Detailed
    - Large
- Challenge
  - How to combine the new data sources with the old ones?

MIT MANAGEMENT SLOAN SCHOOL



MIT MANAGEMENT SLOAN SCHOOL

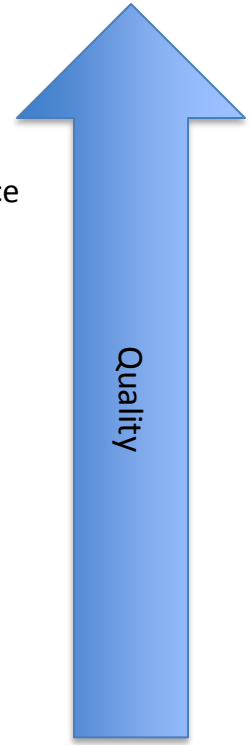
Different Types of Data



MIT MANAGEMENT SLOAN SCHOOL

Organic Data

- Biometric data
- CCTV digital data
- Transactions
  - scanner, credit card, utility usage, health service
- Web pages
  - Scrapping, searches
- Mobile devices
  - GPS-associated data, network development
  - Texts, Blogs, Emails
- Internet of things
  - weather sensors, road sensors, appliances
- Social networks
  - Twitter, Instagram, Facebook



MIT MANAGEMENT SLOAN SCHOOL

Problems in the Data

- Typical Survey Quality Evaluation
  - Representativeness
    - Coverage
    - Sampling
    - Non-response
  - Measurement
    - Validity of procedure (or measure)
    - Measurement error
    - Processing

MIT MANAGEMENT SLOAN SCHOOL

Requirements

Weaknesses that matter

	Survey	Estimation	Forecasting	Measurement
Representativeness				✓
Selection Bias		✓		✓
Reliability and Consistency	✓		✓	✓
Transparency on Data Collection and Treatment	✓	✓		✓
Errors-in-variables		✓	✓	✓
Aspirational (Transactional)	✓	✓		✓
Private (as opposed to public)	✓		✓	✓
Model Uncertainty and Behavioral Changes			✓	✓

MIT MANAGEMENT SLOAN SCHOOL

Measurement

- Official Statistical Offices
  - Federal Statistics
    - GDP, Labor Market Conditions (Unemployment), Inflation, Real Estate, etc.
  - Tremendous challenges
    - Response rates
    - Cost and Budgets
    - Political Manipulation
- Blending organic and designed data
  - Use the big data “good” properties of organic data
  - Use survey data to correct for the biases and the “missing” observations in the big data

MIT MANAGEMENT SLOAN SCHOOL

Online Inflation Rates

Billion Prices Project

MIT MANAGEMENT SLOAN SCHOOL

Billion Prices Project

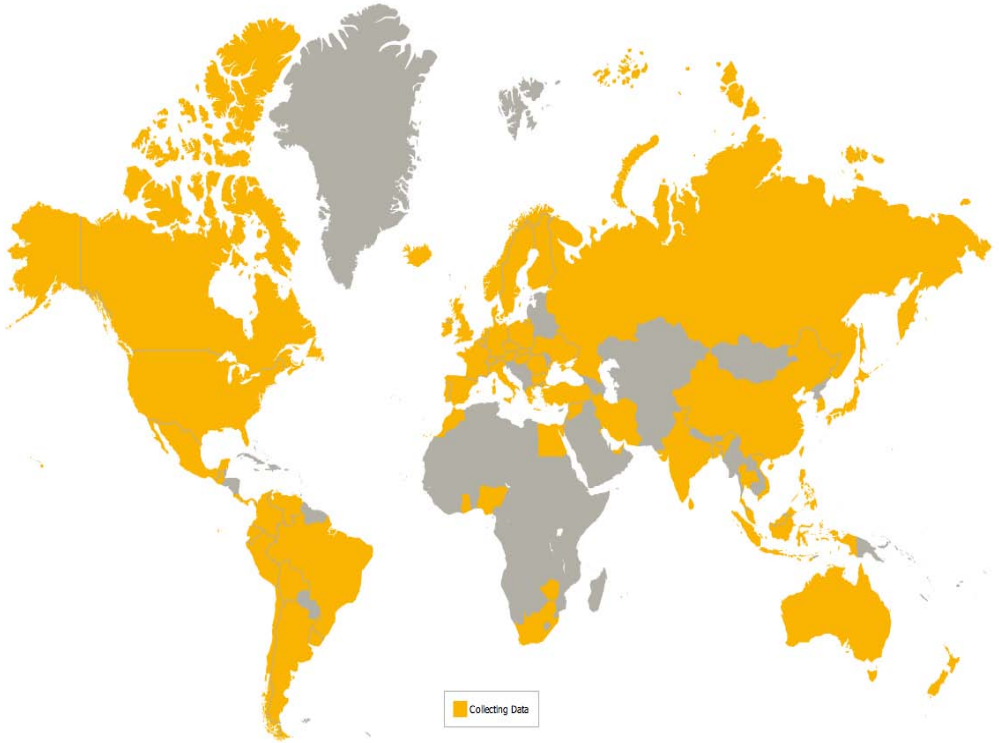
Our Approach to Daily Inflation Statistics

1. *Use scraping technology*
2. *Connect to thousands of online retailers every day*
3. *Find individual items*
4. *Store and process key item information in a database*
5. *Develop daily inflation statistics for ~20 countries*

- *Date*
- *Item*
- *Price*
- *Description*

MIT MANAGEMENT SLOAN SCHOOL

Countries covered

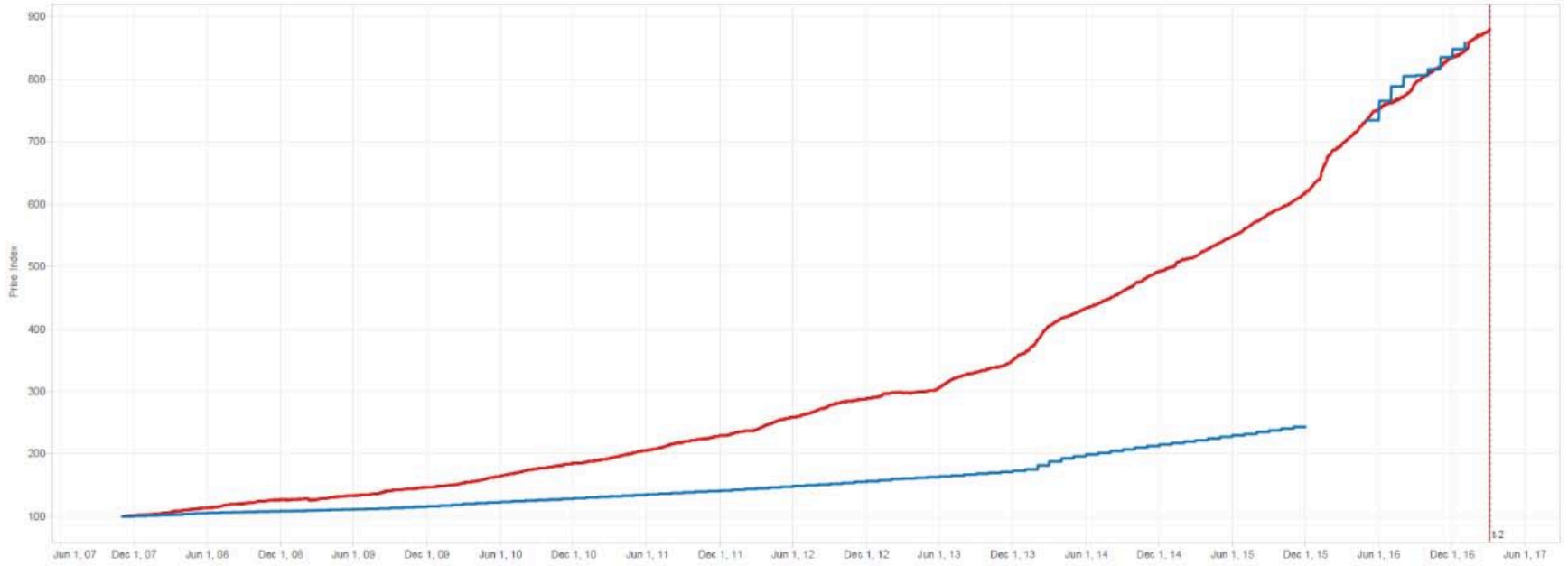


MIT MANAGEMENT SLOAN SCHOOL

USA

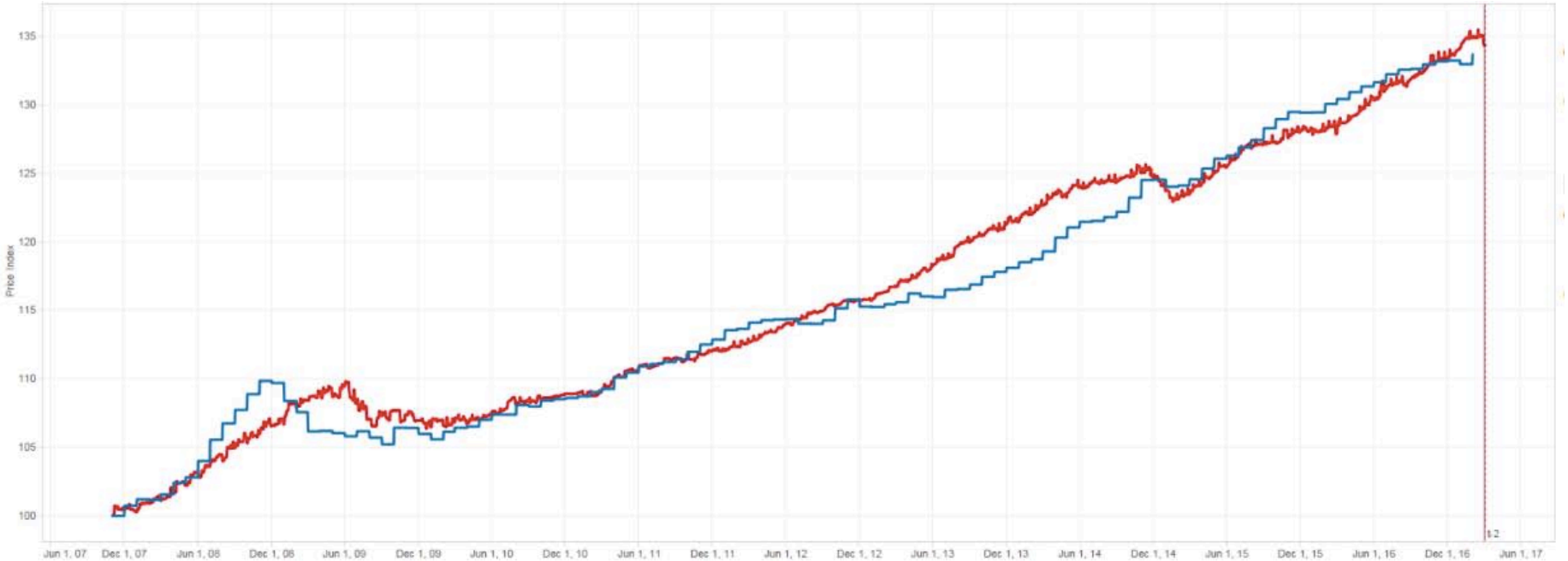
MIT MANAGEMENT SLOAN SCHOOL

Argentina



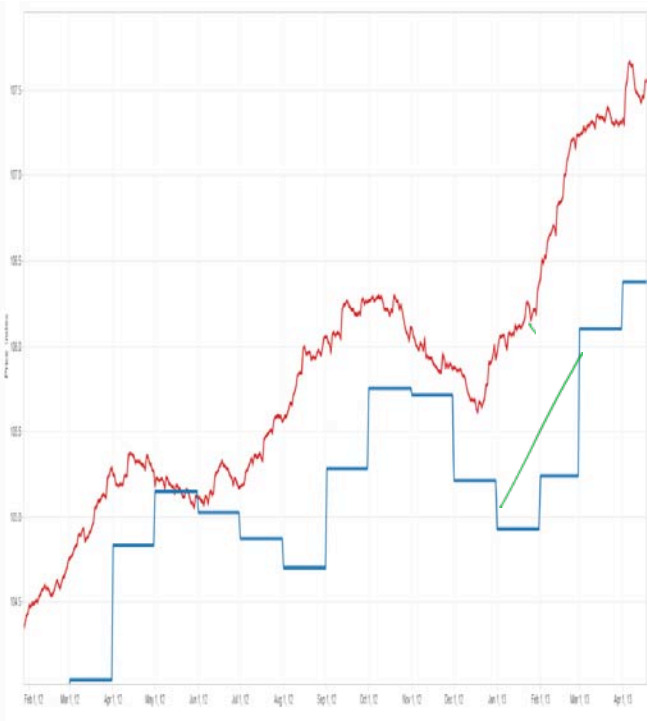
MIT MANAGEMENT SLOAN SCHOOL

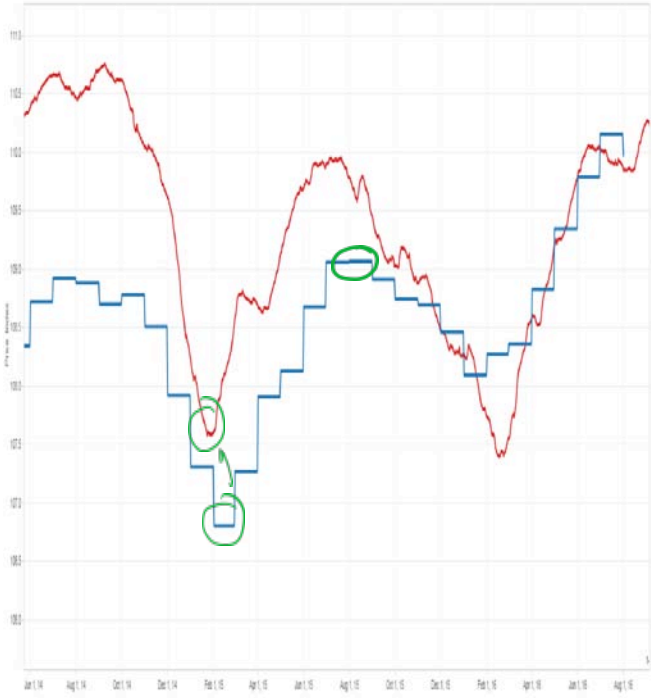
Chile



MIT MANAGEMENT SLOAN SCHOOL

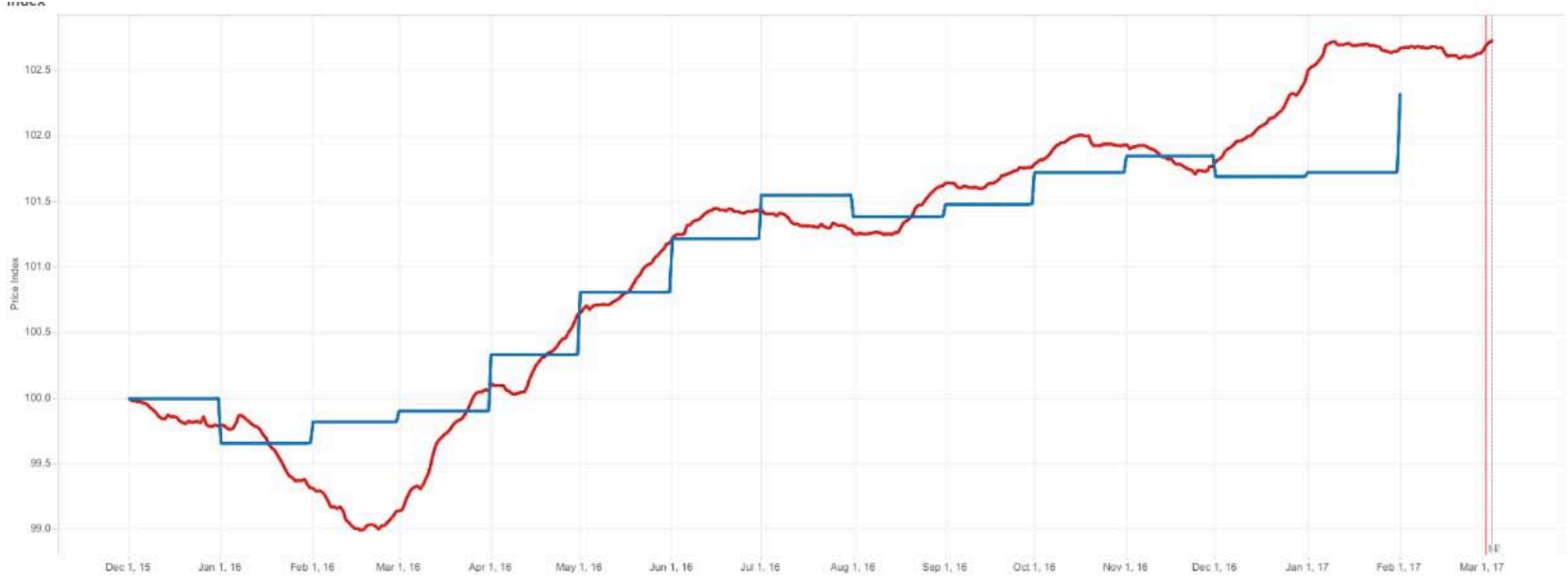
USA: Anticipation





MIT MANAGEMENT SLOAN SCHOOL

Just happened again!



MIT MANAGEMENT SLOAN SCHOOL

In Summary

- Big Data sources (organic data) represents a huge opportunity
  - Understand the weaknesses in the data
- Measurement
  - Statistical offices will not be substituted
  - Blending the data sources is the way to the future

MIT MANAGEMENT SLOAN SCHOOL