Billion Prices Project:
Avoiding the Big Data - Small Info Syndrome

Roberto Rigobon
MIT Sloan, NBER, CSAC
Big distance between Data and Information!

- The world is not lacking of data, it’s lacking of information
  - Need a question before collection
  - Data without purpose, produces problems not information

- Great Data is not a licenses for Bad Econometrics
  - Need to understand phenomena
  - Measuring is not Predicting
  - Correlations are hardly Causal
Inflation and Price Indexes

- Alberto Cavallo

- We want to produce alternative measures of inflation
  - We need prices
  - We need methodology to weight those prices
  - We need to understand product introductions and discontinuations
  - We need to understand store behavior
  - How to collect prices of products not sold on the internet? Services?
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
How do we collect data?

- Our prices are collected from public online sources, using a technique called “web scraping”.

- A software downloads a webpage, analyses the html code, “scrapes” price data, and stores it in a database.
Countries covered
Properties of Inflation Indexes

- **Congruence**
  - Stores keep markups between online and offline prices relatively constant in the medium run

- **Anticipation**
  - Online Price Indexes trends anticipate official inflation shifts.
  - Online prices are easier to change, retailers are more competitive, and consumers have less memory.

- **Demand trends**
  - Changes in inflation trends by retailers reflect changes in the demand they are facing
Argentina
Annual Inflation Rates

Argentina

Australia

China - Supermarket Index

Colombia

Germany

Ireland

Russia

UK

Venezuela
Other Indexes

- Natural Disaster’s Measurement
- Employment
- Real Estate Inflation and Capital Gains
- GDP and Economic Activity
Thailand: When Help Should Be Deployed?

Product Availability In Online Retailers

Bangkok residents hoard food in face of flood
Suttinee Yuvejwattana, Supunnabul Suwannakij, Bloomberg News
Thursday, October 13, 2011

Flood Defenses Are Overrun in Bangkok

River at record high as flood reaches Thai palace
October 29, 2011 | Seth Mydans, New York Times

Bangkok Hopes Flood Barriers Will Hold
By AP / TODD PITMAN AND THANYARAT DOKSONE Saturday, Oct 15, 2011

Notes: Product Availability normalized to 100 on 10/1/2011