

# Google Web Searches as a Dense Survey

Jeffrey D. Oldham

(oldham@google.com)

2015-11-17

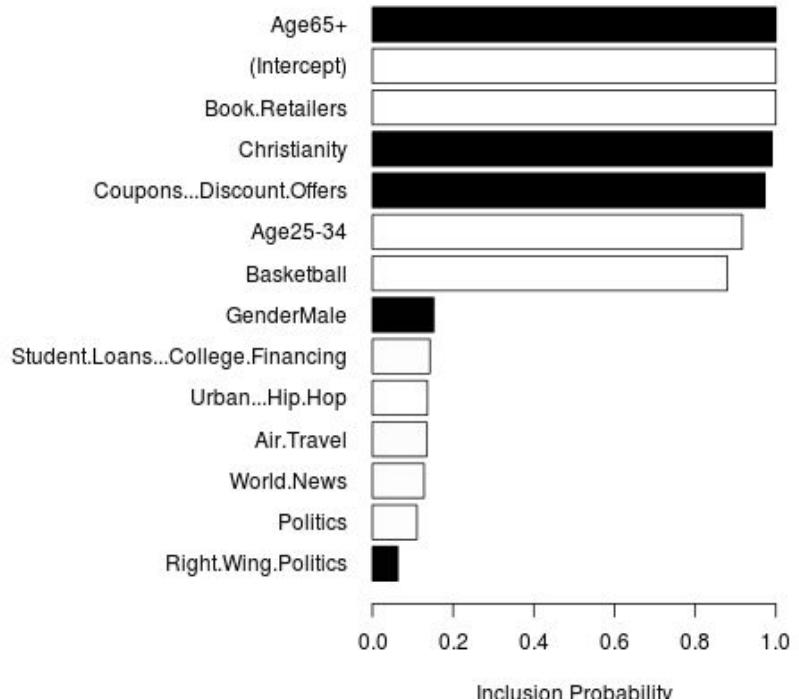
# Survey amplification: Extrapolate geographically

Question: How many people in each state or metro support Obama (in 2012)?

1. Conduct 1000-person national survey:

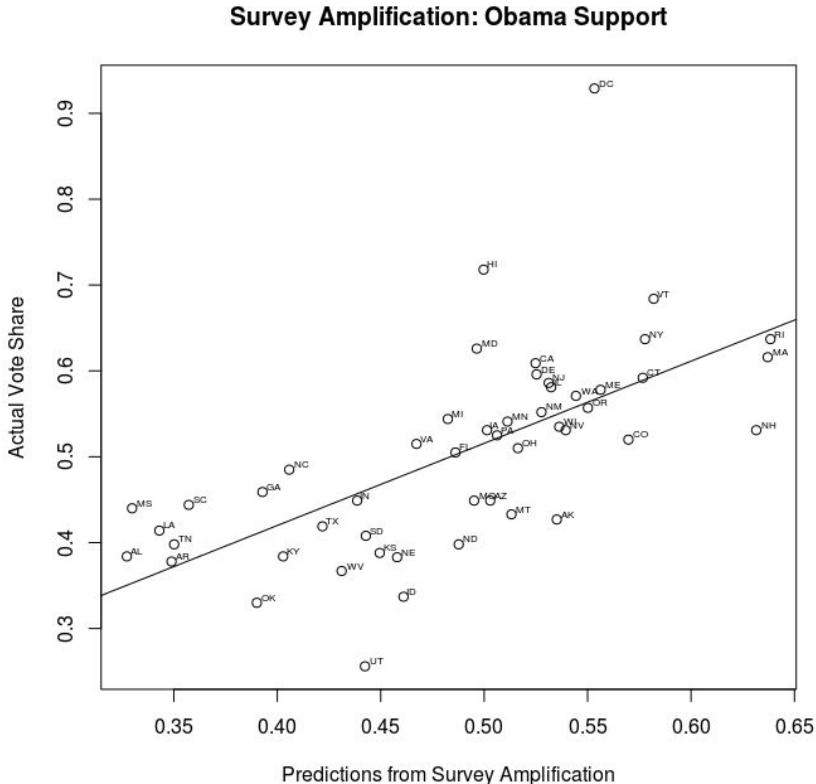
"Do you support Obama in the upcoming 2012 election?"

2. Model at city-level using [Google Trends](#) categories.



# Survey amplification: Predict state (or DMA) support

## A handful of outliers — DC, Hawaii & Utah...



...the majority of **states'**  
**actual vote share**  
**aligns with Survey**  
***Amplified prediction.***

# Google Web Searches = Dense Continuous Survey

Query is a "pushed" indication of interest at time  $t$  in place  $p$ .

- Millions of survey participants.

Aggregate (preserving privacy) by

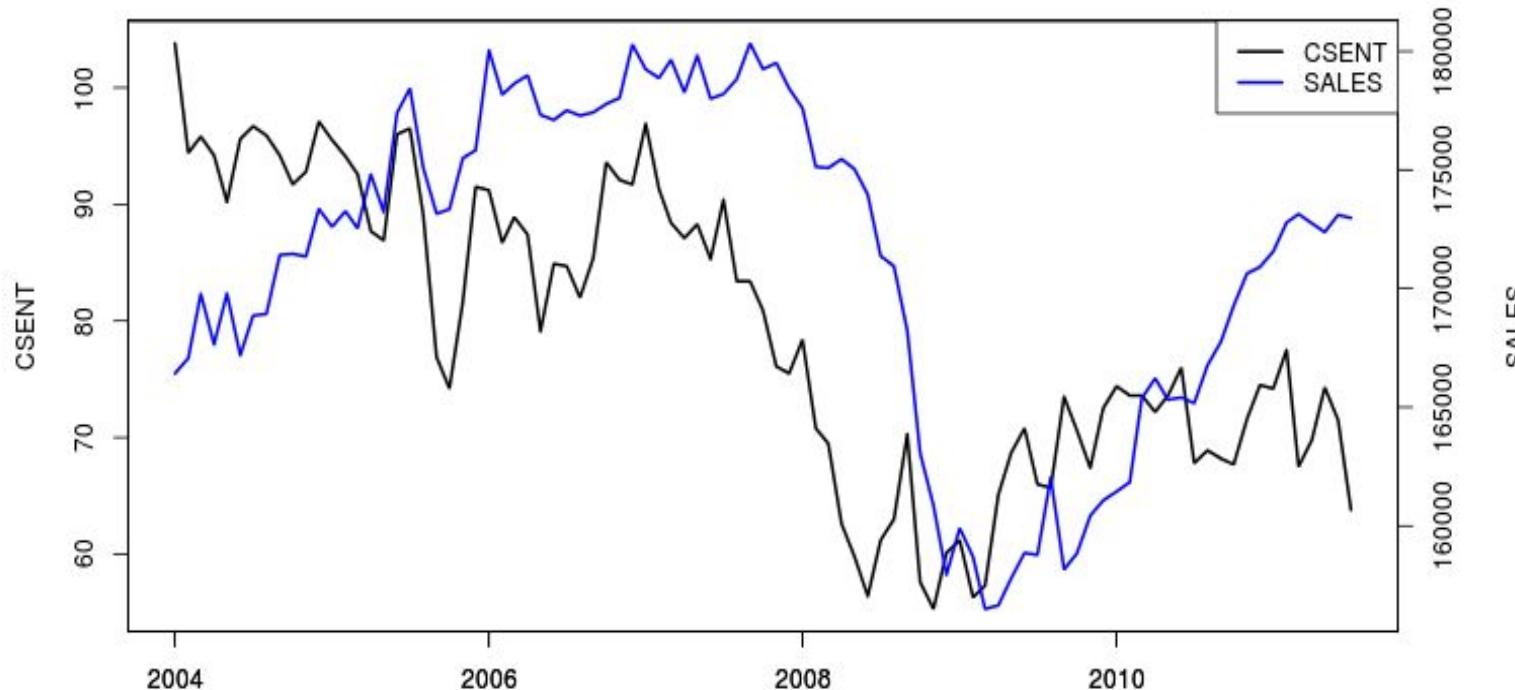
- geography
- time
- interest category.

Survey/geo amplification = small data + post-stratification on geo + category.

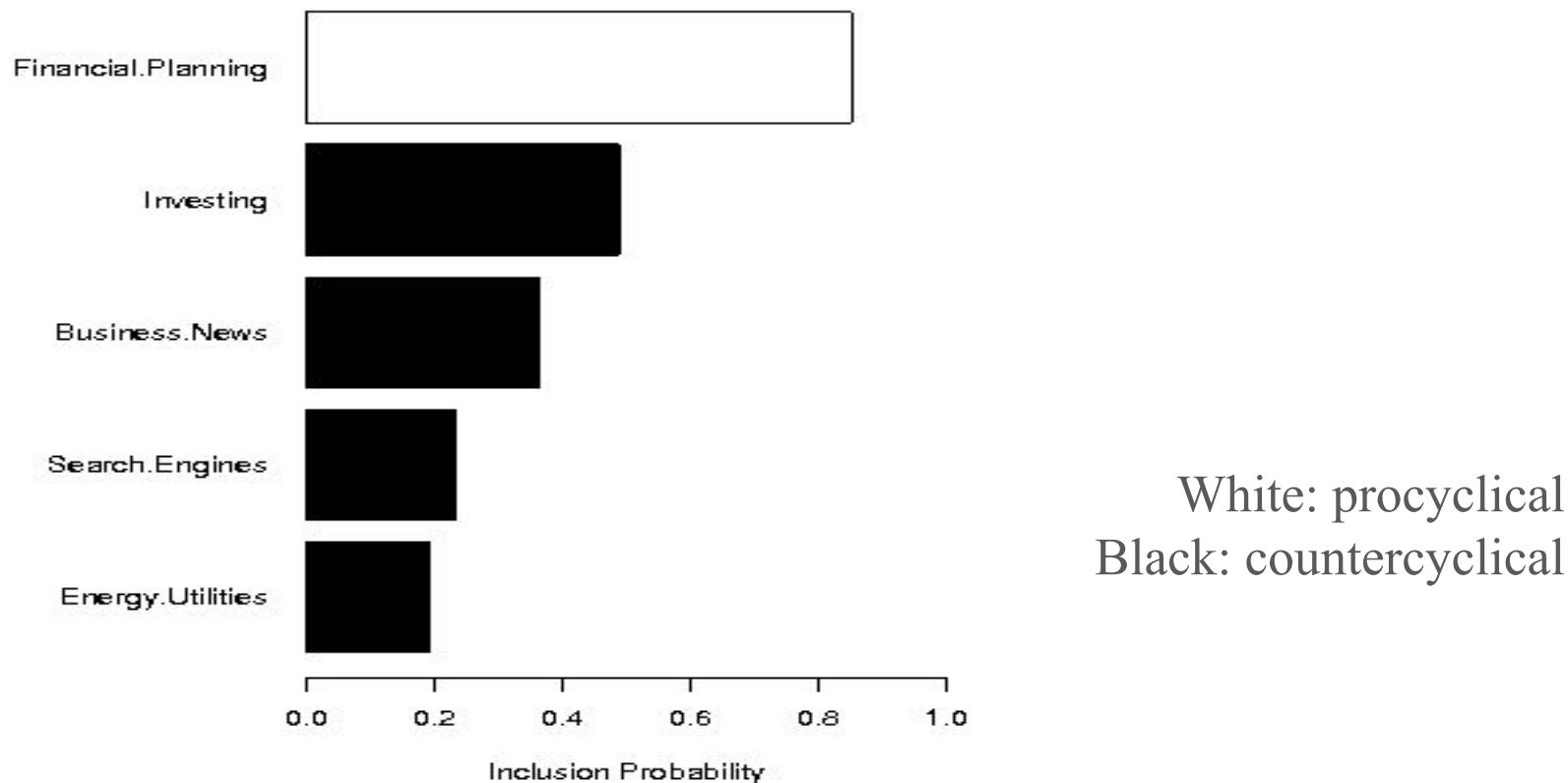
Predicting the present = small data + forecasting on time + category.

# Predicting the present: Real-time statistics

What is consumer sentiment right now?

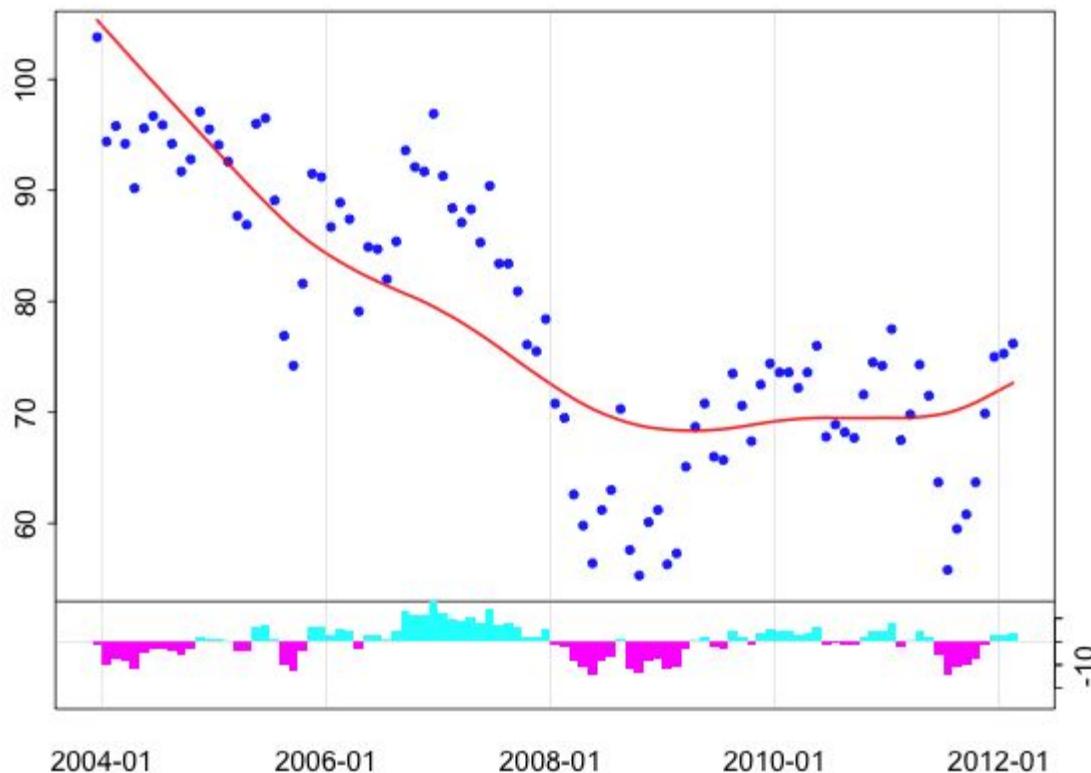


# Model using query categories

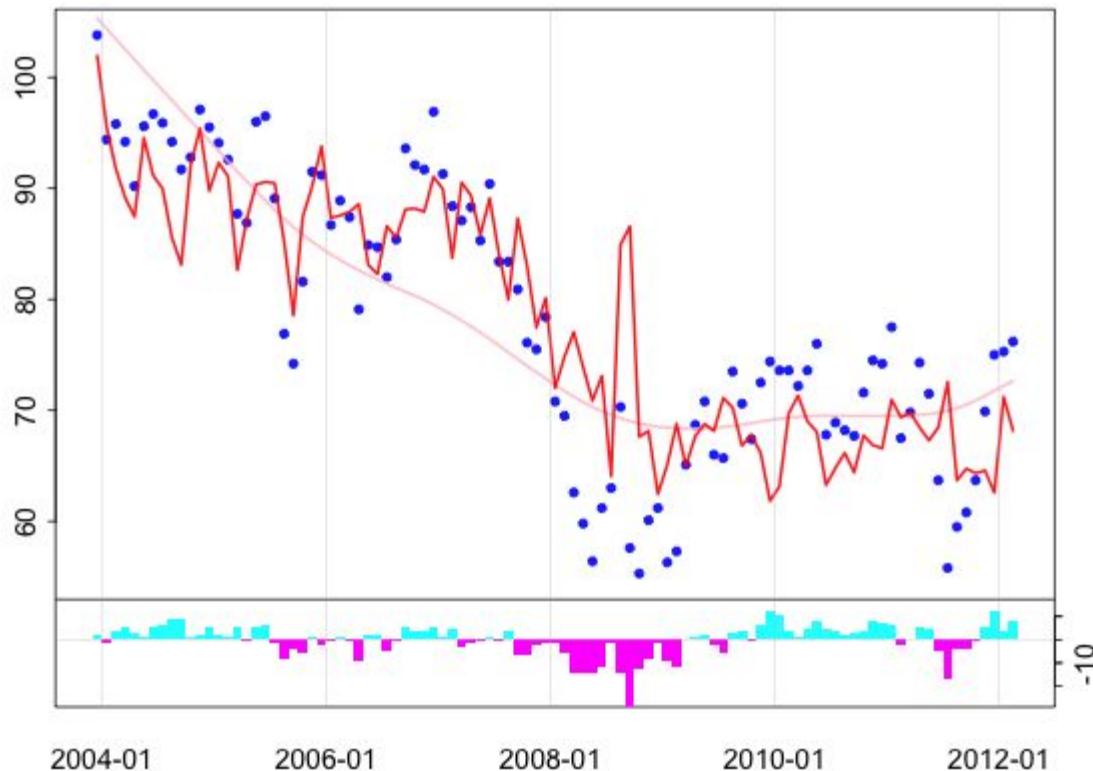


# Trend

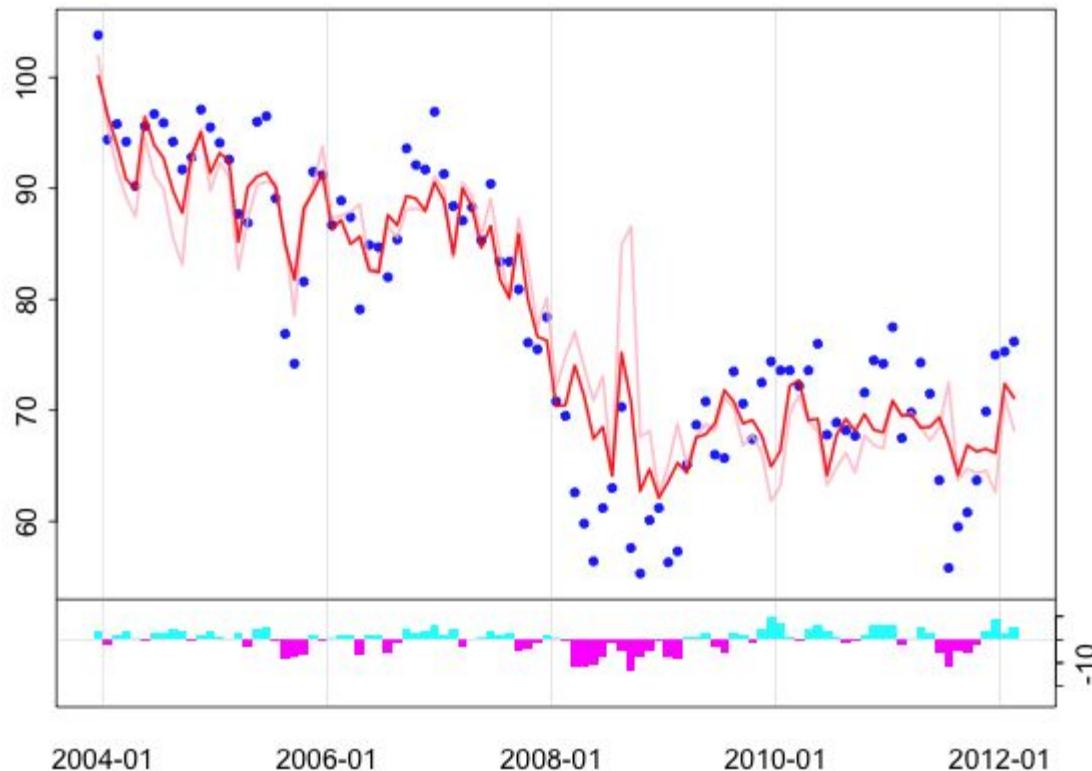
1. trend (mae=5.7134)



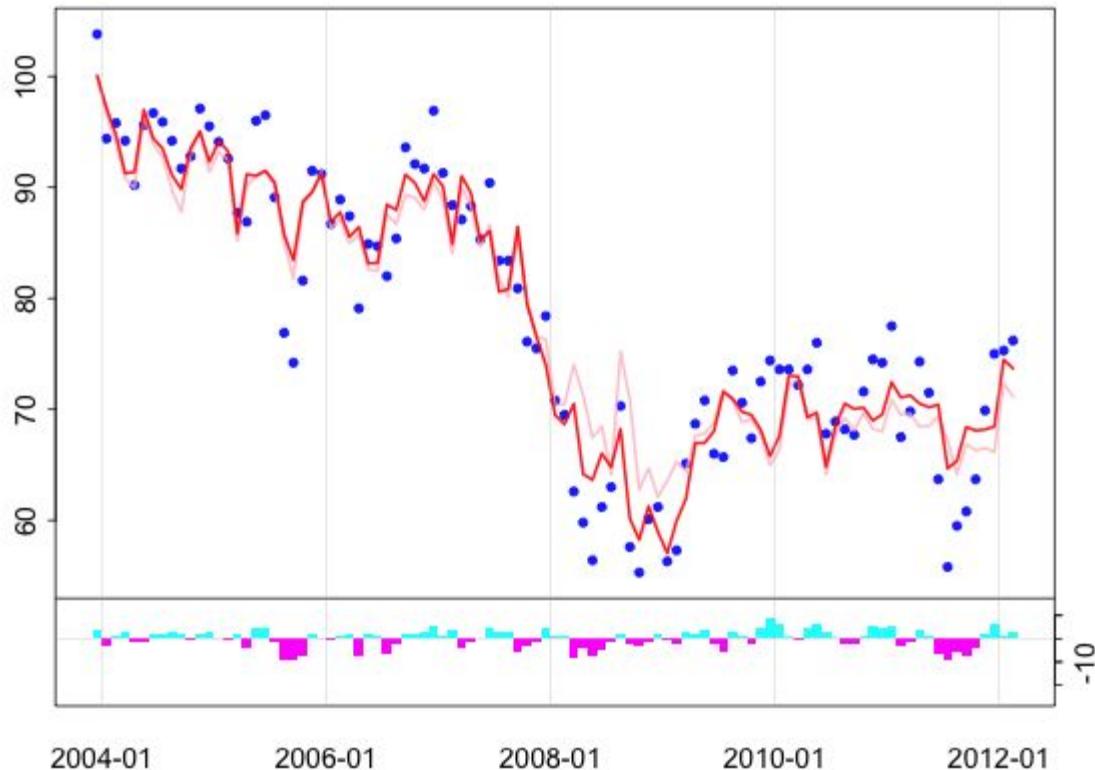
## 2. add Financial.Planning (mae=4.9965)



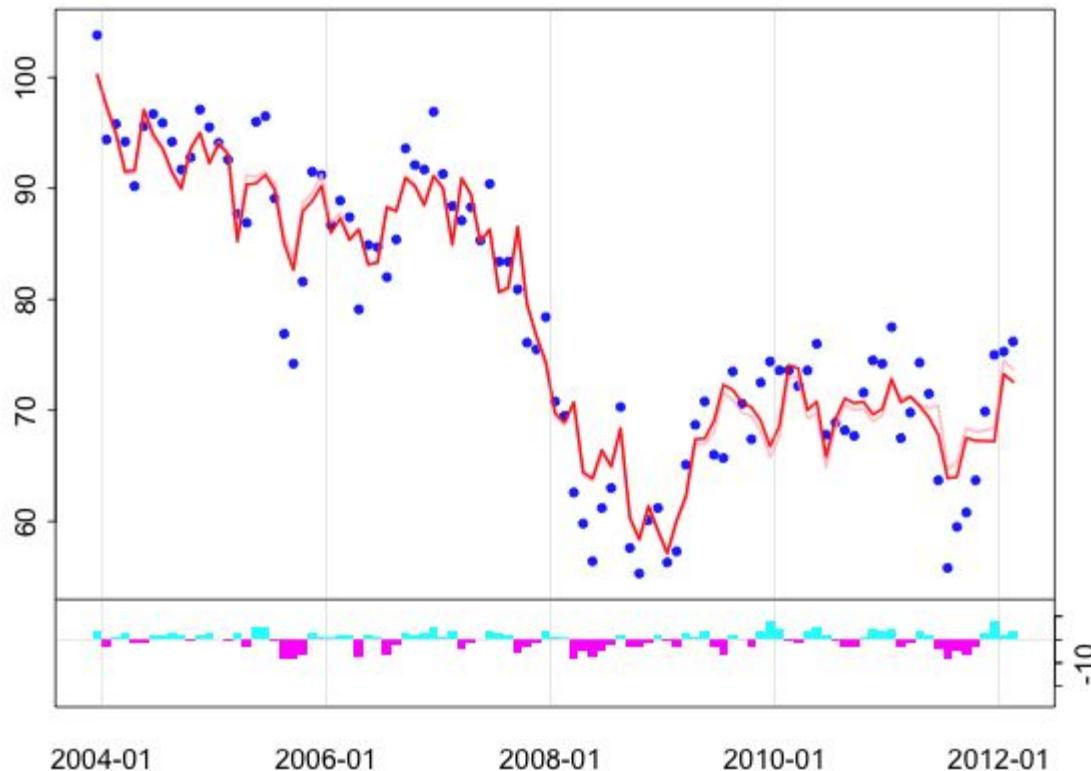
### 3. add Investing (mae=3.8372)



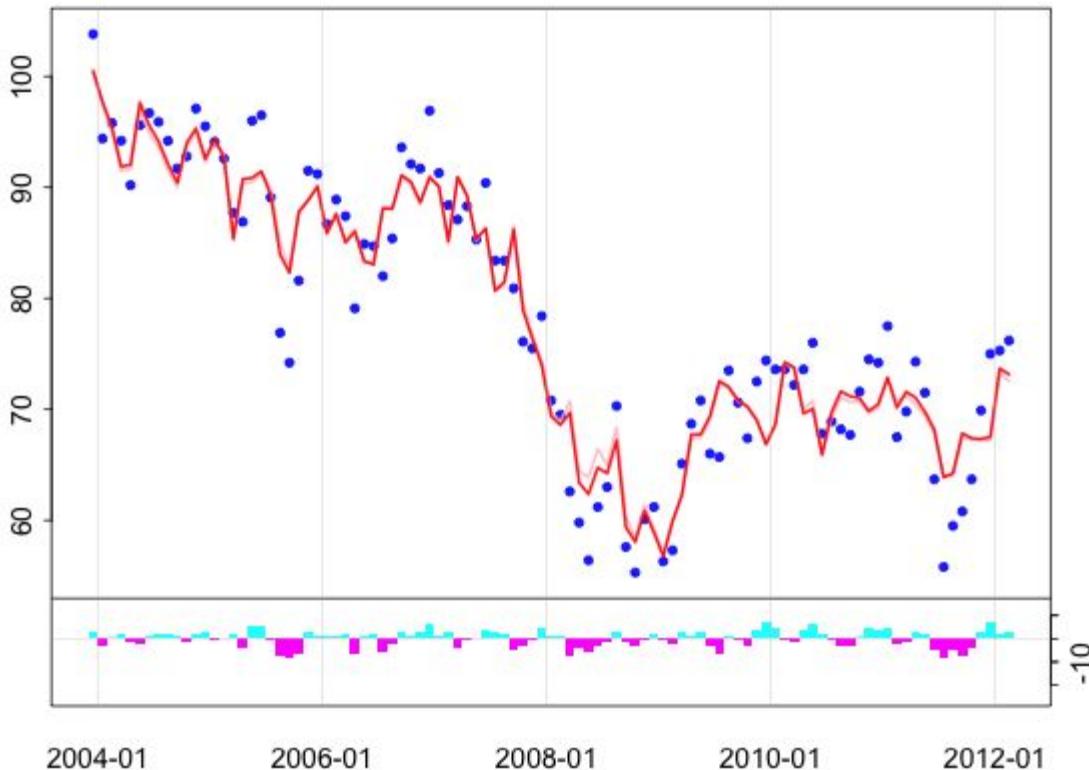
#### 4. add Business.News (mae=3.2226)



## 5. add Search.Engines (mae=3.1455)



## 6. add Energy.Utilities (mae=3.0068)



# Google Web Searches as a Dense Survey

Jeffrey D. Oldham

(oldham@google.com)

2015-11-17