



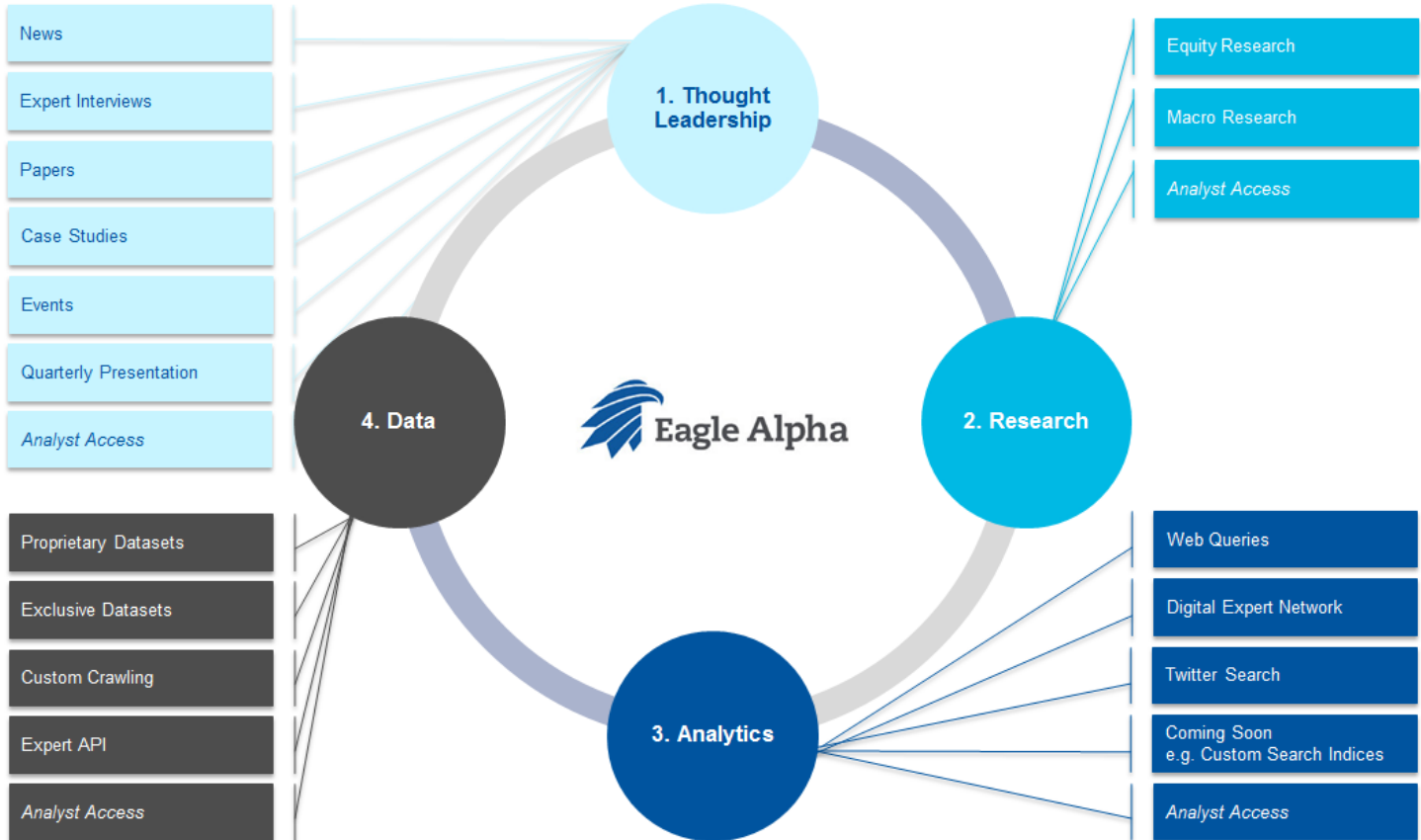
Construction of Google Search Indices by Applying Principal Component Analysis

Sile Li & Shannon Callan, CFA

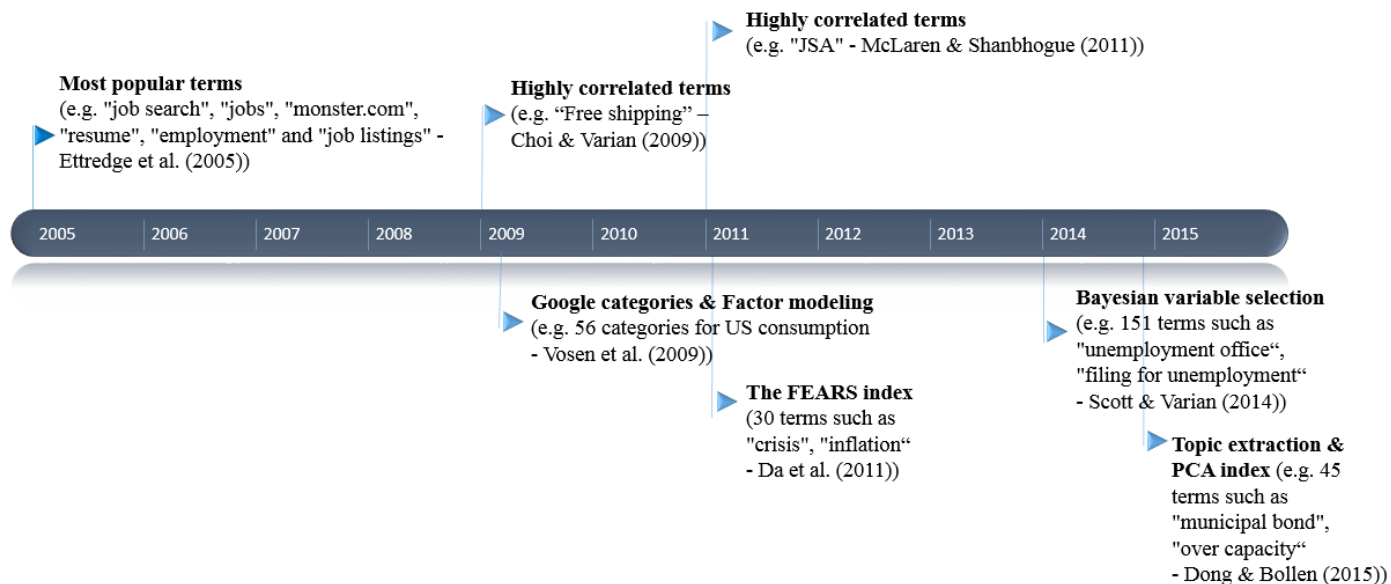
Contacts:
enquiries@eaglealpha.com
(646) 843-6048

20th May 2016

Eagle Alpha Enables Asset Managers to Obtain Alpha from Alternative Data

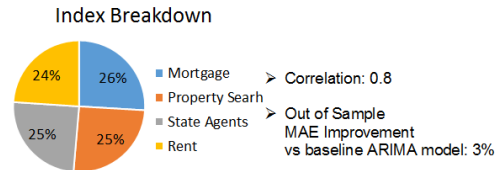
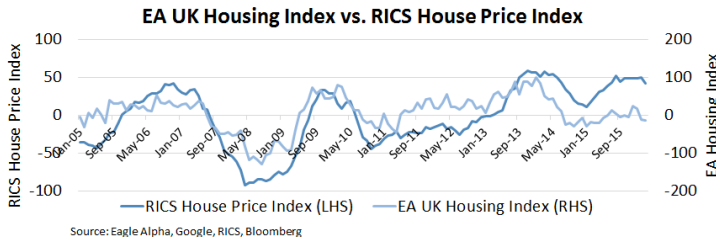
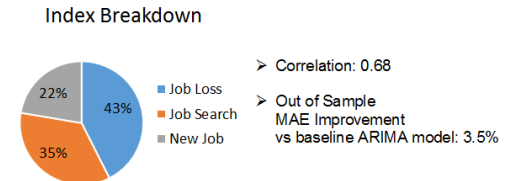
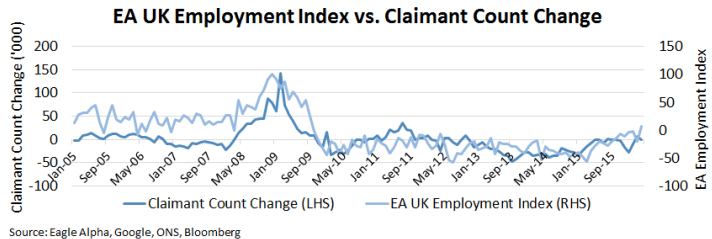
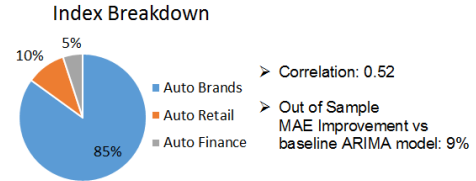
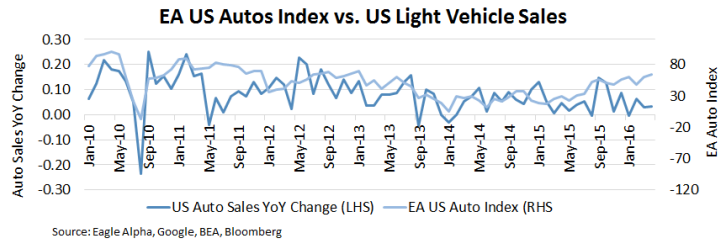
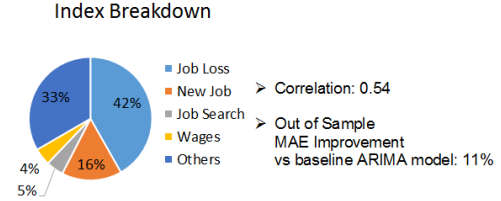
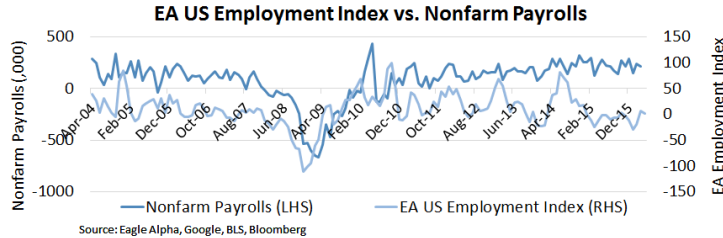


- Internet Search Data (Google, Baidu)
 - ▶ Advantages: new information, continuous information, broadness, easy access
 - ▶ Disadvantages: relatively short history, sampling bias, diverse and dynamic search behaviors
- Studies Have Shown Predictive Value for Economic and Financial Metrics
 - ▶ Unemployment rate, house price, consumption, tourism, inflation, commodity price volatility, stock market return and volatility etc. in various countries.
- Evolution of Methodology



- Term Selection (packages *BMA*, *BSTS*)
 - ▶ Primitive list
 - ▶ Harvard dictionary economic keywords
 - ▶ Find related searches in Google Trends and Google Correlate
 - ▶ Filtering via correlation, Bayesian variable selection, term loading in PCA analysis
- Obtain Data and Process (packages *gtrendsR*, *zoo*, *robustHD*, *seas*)
 - ▶ Convert time frequency, remove outliers, winsorize, and seasonal adjustment
- Apply Principal Component Analysis (function *prcomp*)
- Index Creation
 - ▶ Create index from top principal component(s)
 - ▶ Use of rolling window
 - ▶ Track category loadings within the index

Eagle Alpha's Online Search Index gauges search behaviours related to specific economic activities. It is a composite indicator which measures the co-movement of multidimensional and dynamic search terms.



- Currently: data, reports and analyst access.
- Under Construction: tool for clients.

May 5, 2016
Searches From Those With New Job Gain Most in April
 Overall Index Little Changed

Research: Macro Trends
 Region: United States
 Theme: Employment

Employment Related Searches Little Changed in April
 Eagle Alpha's US Employment Index (see Fig. 1), a measure of online search activity relating to the jobs market, remained virtually unchanged in April (-0.02 points). April's figure was driven primarily by searches relating to the loss of a job (see Fig. 2), although the search category also fell the most in terms of importance on the month. Searches relating to those having gained a new job gained the most in importance in April.

For more information on methodology and a guide on how best to use the data, please see our separate information document, or contact us via macro@eaglealpha.com.

Fig. 1 - EA US Employment Index vs. Nonfarm Payroll

Source: Eagle Alpha, Google, BLS, Bloomberg

Fig. 2 - Eagle Alpha US Employment Index Breakdown

Source: Eagle Alpha, Google

The Eagle Alpha service consists of providing insights from various publicly available sources of information and other third party sources and providing signals of the information to customers who are professional institutional investors acting in the course of business. Content provided to customers through our service (including this document) and the Eagle Alpha website (www.eaglealpha.com) (the "Website") is general in nature and is not tailored to the investment needs of any particular investor. By reading this document, you indicate that you accept the Disclaimer at the end of this report (Appendix 3).

Online Search Tool

http://eaglealpha.com

One Two Three

Region:
 Americas
 US
 Some other country
 Europe
 APAC

Step 2: Index Creation and Analysis

Index Creation Factor Analysis

Set Parameters
 Moving Average:
 3 period
 6 period
 12 period
 Custom

Seasonally Adjust
 Yes
 No

Apply Log
 3 period
 6 period
 12 period
 Custom

Comparative Analysis:
 Compare with existing data
 Employment (US - Nonfarm payroll) Choose File
 Auto Sales (US) No file chosen

Relevant Statistical Output

Metric	Result
Correlation	0.70
Out of Sample Testing vs AR1 (MAE)	5%
Improvement in R2 vs AR1	0.2

Update Download

- Choi, H. and H. R. Varian (2012). Predicting the Present with Google Trends. *Economic Record* 88 (SUPPL.1), 2–9.
- Da, Z., J. Engelberg, and P. Gao (2011). The Sum of All FEARS: Investor Sentiment and Asset Prices. *Review of Financial Studies* 28 (1), 1–40.
- Dong, X. and J. Bollen (2015). Computational Models of Consumer Confidence from Large-Scale Online Attention Data: Crowd-Sourcing Econometrics. *Plos One* 10 (3), e0120039.
- Ettredge, M., J. Gerdes, and G. Karuga (2005). Using Web-based Search Data to Predict Macroeconomic Statistics. *Communications of the ACM* 48 (11), 87–92.
- Hakkio, C., & Keeton, W. (2009). Financial Stress: What is it, How Can it be Measured, and Why Does It Matter? Federal Reserve Bank of Kansas City.
- McLaren, N. and R. Shanbhogue (2011). Using Internet Search Data as Economic Indicators. Technical Report 1, Bank of England.
- Nardo, M., Saisana, M., Saltelli, A., Tarantola, S., Giovannini, E., & Hoffmann, A. (2008). *Handbook on Constructing Composite Indicators: Methodology and User Guide*. OECD.
- Scott, S. L. and H. R. Varian (2014). Bayesian Variable Selection for Nowcasting Economic Time Series. forthcoming in *Economics of Digitization* (July 2012), 1–22.
- Vosen, Torsten Schmidt, S. (2009). Forecasting Private Consumption: Survey-based Indicator VS. Google Trends. *Journal of Forecasting* 578 (January), 565–578.