Predicting Initial Claims for Unemployment Benefits

Hyunyoung Choi, Hal Varian

July 5, 2009

What would you search for if you thought you might lose your job? Typical searches might be queries like [unemployment office], [apply for unemployment], [jobs], and so on.

Google Search Insights and Google Trends classifies search queries like these into two categories, 'Local/Jobs' and 'Society/Social Services/Welfare & Unemployment'. It is tempting to think that Google searches in these topics may be related to filings for unemployment benefits.

In our earlier work([Choi and Varian(2009)]), we described how Google Trends data could help predict economic time series of interest. Here we apply the methodology developed in that paper to U.S. unemployment time series.

The "Initial Jobless Claims" is a weekly report issued by US Department of Labor.[i] It tracks the number of people filed for the unemployment benefits and it is considered a leading indicator of labor market. The claims are reported at the national and state level and the data data is available from first week of 1987. Advance initial claims weekly reports at the national level data are released 5 days after the week ends.

Table 1 shows a sample of release schedule for June 2009 at national level along Initial Claims and Google Trends data. When the data is released at 7/2/09, Google Trends data is available to the week of release, which means Google Trends is 7 days ahead of government's release schedule.

Week	5/31 - 6/6	6/7 - 6/13	6/14 - 6/20	6/21 - 6/27	6/28 - 7/4
Initial Claims	605K	612K	630K	614K	Release
Continued Claims	6.71M	6.76M	6.70M		at 7/9/09
Jobs	3%	1%	0%	-2%	-4%
Welfare & Unemployment	38%	36%	40%	44%	40%

Table 1: Initial Claims and Google Trends Schedule from 5/31/09 to 7/4/09

[i] http://www.dol.gov/opa/media/press/eta/ui/current.htm

Figure 1(a) and 1(b) depict Initial Claims vs. two labor market related categories in Google Trends from 2004. Initial Claims are plotted against Google Trends to compare the trends over time. These three time series show the similar cyclical pattern and up trends since 2008.

A committee of the National Bureau of Economic Research has declared that US has been in recession since December 2007.[ii.] We note that initial claims for unemployment benefits have been increasing since mid 2007(Figure 1). We could utilize the time series since 2004 under the assumption of no structural change for the past 5 years. Or we could use the data since December 2007 under the assumption that there is structural difference since recession started and there is consistent seasonality in labor market.

Initial claims have a very good record as a leading indicator. Macroeconomist Robert Gordon points out that there is a "surprisingly tight historical relationship in past US recessions between the cyclical peak in new claims for unemployment insurance (measured as a four-week moving average) and the subsequent NBER trough."[iii.] Other economists, such as James Hamilton, have also documented this relationship.[iv.]

Denote y_t, Jobs$_t$ and Welfare$_t$ as the time series of Initial Claims, 'Jobs' and 'Welfare & Unemployment' categories in Google Trends respectively. We applied standard ARIMA model selection procedures and selected AR(1) model as a baseline. Then we added the Google Trends series to see how much this improved prediction. This procedure follows the logic described in *"Predicting the Present with Google Trends."*[v.] The baseline and alternative model with Google Trends data for the Long term (all data) and Short Term (recession only) model are as follows:

$$\text{Baseline Model} \quad : \quad \log(y_t) = \text{Intercept} + \phi \log(y_{t-1}) + e_t$$

$$\text{Alternative Model} \quad : \quad \log(y_t) = \text{Intercept} + \alpha \text{Jobs}_t + \beta \text{Welfare}_t + \phi \log(y_{t-1}) + e_t$$

Table 2 summarizes the model fit. There is positive correlation between initial claims and the search related to 'Jobs' and 'Welfare & Unemployment'. Even though the time series plot suggest a change in Initial Claims after 2008, the model predicts better with the longer time series. With the Google Trends series, the model fit is improved significantly and the out-of-sample mean-absolute-error estimated with the rolling window for the past 24 weeks is decreased by 15.74% for the long term model and 12.90% for the short term model.

The national level initial claims are released up to the week of 6/21/09-6/27/09 and the figure is 613K, unexpectedly smaller than the consensus, 619K. Decrease in Google Trends
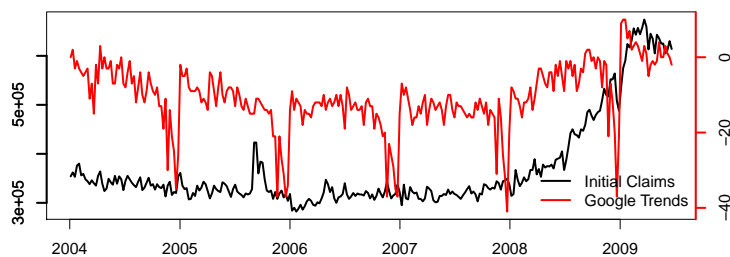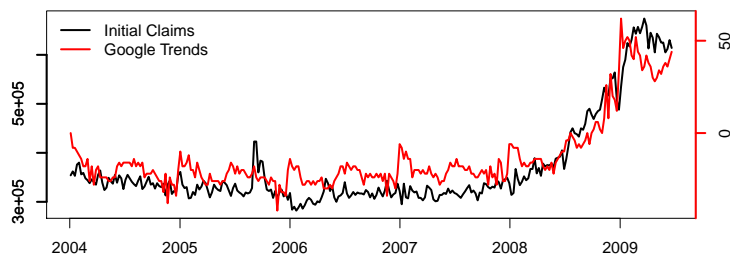

[ii.] http://www.nber.org/cycles/recessions.html
[iii.] http://www.voxeu.org/index.php?q=node/3524
[iv.] http://www.econbrowser.com/archives/2009/04/another_green_s.html
[v.] http://google.com/googleblogs/pdfs/google_predicting_the_present.pdf

(a) Jobs

(b) Welfare & Unemployment

Figure 1: Initial Claims vs. Labor Market related Google Trends

implies decrease in Initial claims. The Google Trends index for the weeks ending 7/4/09 were down and this could indicate a continuous decrease in Initial Claims for the coming weeks.

Summary

We have shown how Google Trends data can help predict initial claims in initial claims for unemployment benefits in the United States. [Askitas and Zimmerman(2009)] and [Suhoy(2009)] have examined similar unemployment data for Germand and Israel, respectively, and also found significant improvements in forecasting accuracy by using Google Trends. We hope that other researchers will continue to examine these data.

|		Baseline Model				Alternative Model					
		Intercept	ϕ	σ	MAE	Intercept	ϕ	Jobs	Welfare	σ	MAE
LT	Est	0.1269	0.9902	0.0443	3.24%	1.6498	0.8727		0.0014	0.0429	2.73%
	SE	0.1618	0.0126			0.3754	0.029		0.0003		
ST	Est	0.2174	0.9839	0.0432	3.10%	1.792	0.8632	0.0014	0.0010	0.0398	2.70%
	SE	0.2632	0.0202			0.5541	0.0427	0.0006	0.0004		

Table 2: Summary of Long Term Models and Short Term Models

References

[1] Nikos Askitas and Kalus F. Zimmerman. Google econometrics and unemployment forecasting. Applied Economics Quarterly, 55:107.120, 2009
URL: `http://ftp.iza.org/dp4201.pdf`.

[2] Hyunyoung Choi and Hal Varian. Predicting the present with google trends. Technical report, Google, 2009
URL: `http://google.com/googleblogs/pdfs/google\_predicting\_the\_present.pdf`.

[3] Tanya Suhoy. Query indices and a 2008 downturn: Israeli data. Technical report, Bank of Israel, 2009.4. URL: `http://www.bankisrael.gov.il/press/eng/090709/090709e.htm`
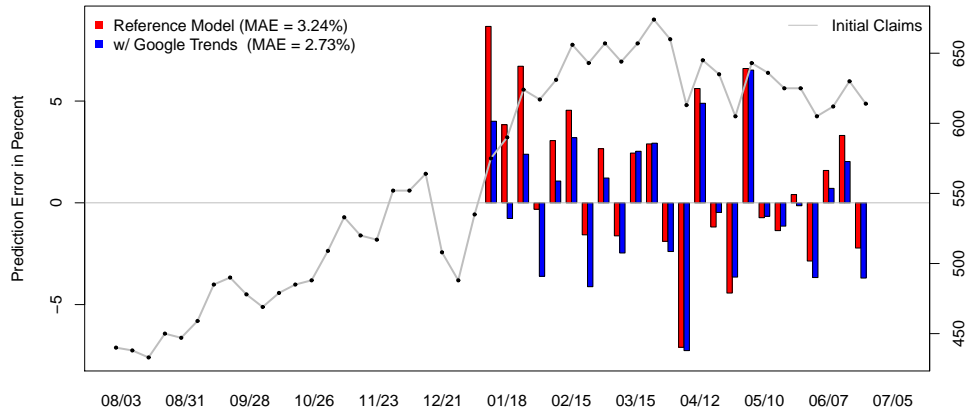
(a) Long Term Model
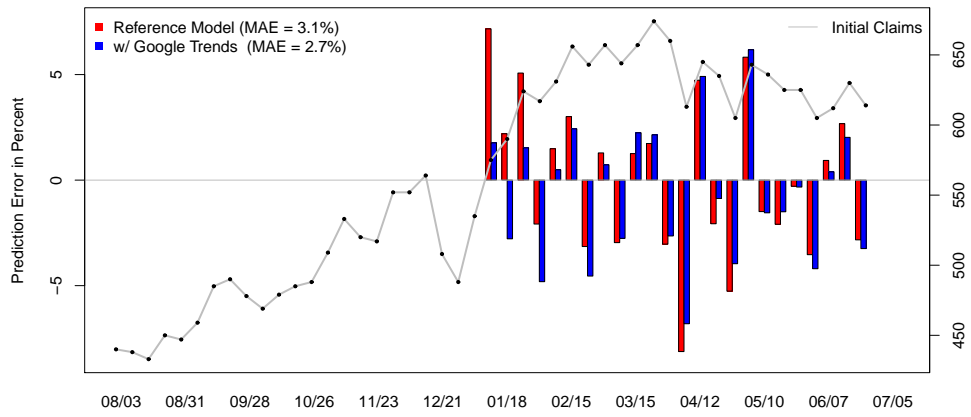
(b) Short Term Model

Figure 2: Prediction MAE from LT and ST Model