



Center for
Research
On the
Wisconsin
Economy
University of Wisconsin- Madison

Forecasting Initial Unemployment Claims using Google Searches

Noah Williams

Center for Research on the Wisconsin Economy, UW-Madison

March 23, 2020

Overview

In the wake of the unprecedented economic disruptions caused by COVID-19, there is an increased need for high-frequency data. The economies in many locations around the globe have ground to a standstill, as businesses have shut down and workers have been ordered to shelter in place in order to try to slow the spread of the virus. In the United States in particular, these shut-downs have led to a rapid increase in layoffs, which will eventually show up in official economic statistics as record-high job losses and increases in unemployment rates. But the official data are produced with a lag, providing a retrospective look at a rapidly-evolving economic situation. In this brief I show that a relatively accurate picture of current economic conditions (at least based on previous experience) can be found at a high frequency by using Google search data. Moreover while this note focuses on the United States as a whole, the same approach can be applied at a more disaggregated level to study individual states.

In particular, I show that weekly Google search trends related to unemployment provide an accurate forecast of initial unemployment claims. **For the current week (ending March 21), I forecast that US initial unemployment claims will top 1.6 million**, which is roughly twice the previous record high.

Forecasting Initial Unemployment Claims using Google Searches

Initial unemployment claims are one of the most widely watched leading indicators, as their changes provide an important signal on labor market conditions. Moreover, since the labor market separations in the COVID-19 pandemic are driven by layoffs (rather than quits), initial claims provide an especially strong signal now about coming unemployment. While some [states](#) have provided preliminary daily data on initial claims, official data is provided weekly and only released with a lag. My forecasts with Google search trends can be updated daily, and as noted can also produce state-level forecasts (which will be forthcoming). In future work I will analyze in more detail the relation between initial claims and unemployment, to show how the projected spike in initial claims will affect future unemployment rates.

Data and Background

A [number](#) of [previous papers](#) have used [Google search](#) trends in different countries around the world to forecast unemployment rates. In general, these papers show that Google searches for different employment- or unemployment-related search terms add significant power for unemployment forecasts. However, especially in the current COVID-19 crisis, I am more interested in tracking rapid labor market changes – increases in unemployment rather than levels of unemployment – and thus I focus on initial unemployment claims. Moreover, this narrower objective allows me to focus on a more specific search trend related to filing unemployment claims (“unemployment office”) rather than a broader set of searches (such as “jobs”, “laid off”, or “unemployment”) as the previous literature has considered.

In particular, I take as the variable of interest the weekly initial unemployment claims series produced by the US Department of Labor, not seasonally adjusted (ICNSA on FRED). The explanatory variable is the Google search trend for the term “unemployment office.” As noted, related but broader measures have been used in previous studies. But after some experimentation I found that this particular search seemed to work best. For example, the broader search topic “unemployment benefits” followed the same broad trends, but had a spike in 2014 when the federal extended benefits program expired. Google provides weekly data back to January 2004, but only allows users to download 5 years at a time, and each span is indexed to peak at 100. To link together the full time series, I downloaded separate 5 year spans with overlapping observations to adjust the scale and re-index. I scale the index so that it matches the mean in initial claims.

The two time series, from 1/1/04 through 3/14/20 are shown in Figure 1. The figure shows that two series are very highly correlated, with a correlation coefficient of 0.82. Further, the series seem to display not only the same broad trends, but very similar cyclical spikes.

Forecasting Initial Unemployment Claims using Google Searches

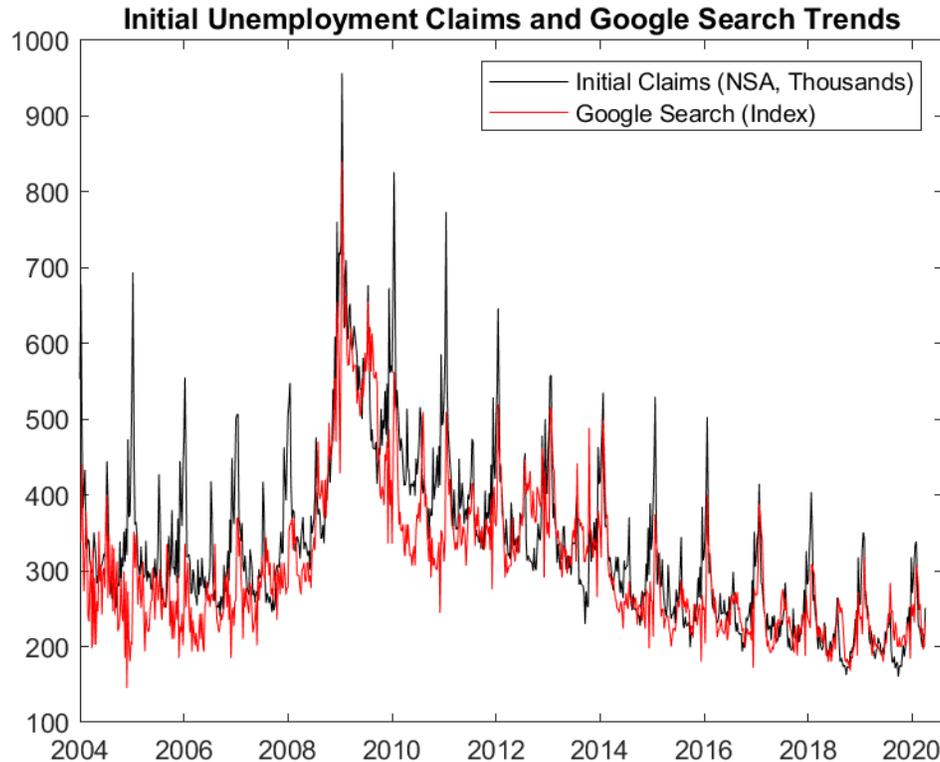


Figure 1: Weekly initial unemployment claims and Google search trend index for “unemployment office,” from January 1, 2004 through March 14, 2020.

Forecasting Initial Claims

I now turn to forecasting initial claims based on the search index. Perhaps more accurately, this is “nowcasting” as I use the data from the same week. That is, the search trend (dated on a Sunday) for the week *beginning* March 15, 2020 is used to predict the initial claims (dated on a Saturday) for the week *ending* March 21, 2020. As of the time of this writing (3/23/20), the search trend is available but the initial claims are not.

I experimented with several different time series VAR specifications, but found that a simple specification with no lags and only the contemporaneous response was best able to capture the periods of sharp increases in initial claims. To deal with the clear seasonality in claims (as well as searches), I included week-of-year dummies as well in each specification. This simple model fit quite well, with an adjusted R-squared of 0.82 and a root mean square error of 48 (thousand).

Moreover, as shown in Figure 2, the predicted values closely match the actual data, particularly in periods around the 2008 recession. The left panel of the figure shows the period from mid-2007 through early 2010. As is now well known, the rise in unemployment in the early part of the 2008 recession was due mostly to a slowdown in hiring rather than an increase in layoffs. This is shown by initial claims remaining

Forecasting Initial Unemployment Claims using Google Searches

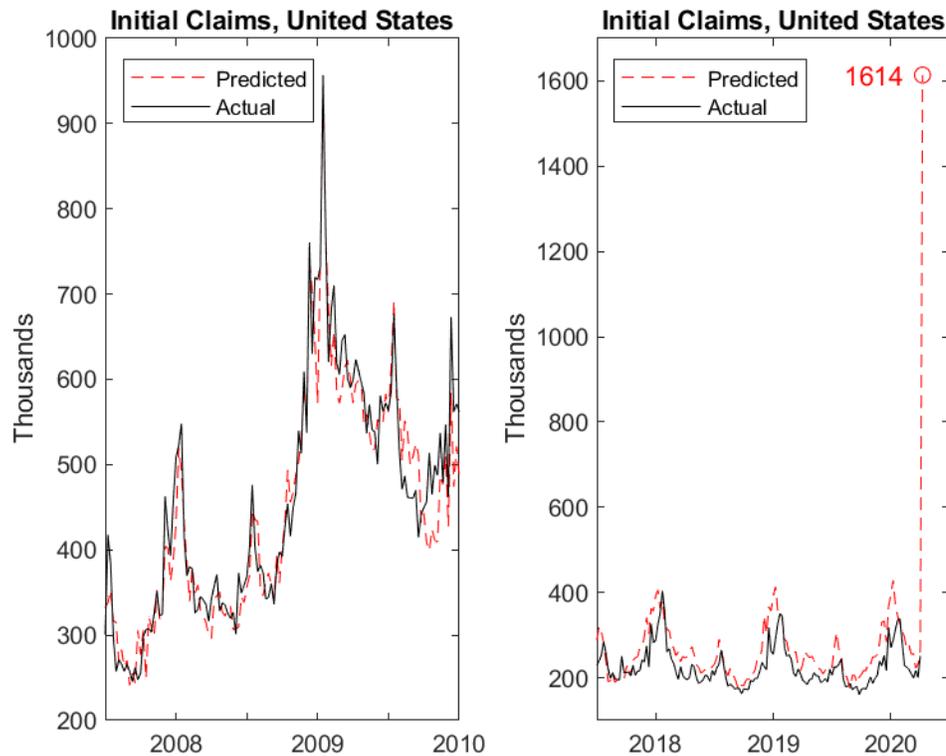


Figure 2: Predicted and actual weekly initial unemployment claims in two separate time periods: around the 2008 recession and recently.

relatively low throughout most of 2008, before increasing rapidly in 2009. The predictions from the Google search trends match quite well the spike in initial claims.

The right panel shows the recent predictions, including the last data point of the week ending March 21, 2020. While the early part of 2020 had been characterized by historically low levels of initial unemployment claims, the mass shutdowns and restrictions of economic activity have led to mass layoffs. Correspondingly, there was unprecedented search activity for “unemployment office” increasing over the course of that week by a factor of roughly 7. While caution should certainly be advised in extrapolating so far beyond past experience (and the relevant model sample), the model predicts that initial unemployment claims for the week will total over 1.6 million. This would roughly double the previous record high.

In future work, I will look at state level data, as the variation in the spread of COVID-19 and the governmental responses have affected states differentially. Moreover, some of the states have released advanced data which can also be used to validate the model. Further, I plan to estimate the size of the increase in unemployment rates that will result from this spike in initial claims.