Google It Up! A Google Trends-based Uncertainty index for the United States and Australia

Efrem Castelnuovo\textsuperscript{a,b,*}, Trung Duc Tran\textsuperscript{a}

\textsuperscript{a} University of Melbourne, Australia
\textsuperscript{b} University of Padova, Italy

HIGHLIGHTS

* We develop uncertainty indices for the United States and Australia based on freely accessible, real time Google Trends data.
* Our Google Trends Uncertainty (GTU) indices correlate positively with a variety of existing proxies for uncertainty.
* VAR investigations point to an economically and statistically significant contribution by GTU shocks to unemployment dynamics in the United States.
* The contribution of GTU shocks to unemployment dynamics in Australia is found to be mild.
* Monetary policy shocks in Australia are a more important driver of unemployment than uncertainty shocks.

ABSTRACT

We develop uncertainty indices for the United States and Australia based on freely accessible, real time Google Trends data. Our Google Trends Uncertainty (GTU) indices are found to be positively correlated to a variety of alternative proxies for uncertainty available for these two countries. VAR investigations document an economically and statistically significant contribution to unemployment dynamics by GTU shocks in the United States. In contrast, the contribution of GTU shocks to unemployment dynamics in Australia is found to be much milder and substantially lower than that of monetary policy shocks.

© 2017 Elsevier B.V. All rights reserved.

1. Introduction

"While there has been substantial progress, a range of questions remain open around the measurement, cause, and effect of uncertainty, making this a fertile area of research." (Nicholas Bloom, Stanford University, \textit{Journal of Economic Perspectives}, 28(2), \textit{Bloom, 2014}, p. 154)

\textsuperscript{⋆} We thank Pierre-Daniel Sarte (Editor), an anonymous referee, Giovanni Caggiano and Giovanni Pellegrino for useful comments. Financial support by the Australian Research Council via the Discovery Grant DP160102281 is gratefully acknowledged.

\textsuperscript{*} Correspondence to: Melbourne Institute and Department of Economics, University of Melbourne, 111 Barry St., FBE building, level 5, 3010 Melbourne (VIC), Australia.

E-mail address: efrem.castelnuovo@unimelb.edu.au (E. Castelnuovo).

This paper constructs Google Trends-based uncertainty indices (GTU indices henceforth) for the United States and Australia. These indices are based on uncertainty-related keywords frequently mentioned in reference economic documents like the Federal Reserve’s Beige Book for the United States and the Reserve Bank’s Monetary Policy Statement for Australia. These documents gather information on current economic conditions based on interviews with key business contacts, economists, and market experts (among other sources). Hence, they are likely to be a good proxy of entrepreneurs’ uncertainty as regards future business conditions.

The choice of developing an uncertainty measure for the United States enables us to validate our novel index by comparing it with a large number of alternative proxies for uncertainty recently proposed by the literature. Moreover, the U.S. economy experienced a severe recession in 2007–09, and it hit the zero lower bound (ZLB) in December 2008, where it remained for seven years. Recent contributions document that recessions and the ZLB may have
Indeed, magnified the negative real effects of uncertainty shocks in the United States (see Nodari, 2014; Caggiano et al., 2014; Caggiano et al., 2017a for the role played by recessions, and Basu and Bundick 2017; Caggiano et al., 2017b for that of the ZLB). Differently, Australia is the only industrialized country belonging to the G10 which has not experienced a recession since 1991, and its policy rate has never hit the zero lower bound in recent times. Hence, it represents an interesting laboratory to quantify the real effects of uncertainty shocks – identified with unexpected changes of our novel GTU index – and contrast them with the U.S.-related results.

We find our GTU measures to be positively correlated with existing measures of uncertainty. According to our VAR investigations, GTU shocks are significant drivers of unemployment and prices in the United States, with a contribution larger than that provided by monetary policy shocks. By contrast, GTU shocks play a minor role as drivers of the Australian unemployment. The difference in these results possibly highlights the role played by the 2007–09 great recession and the ZLB in magnifying the effects of uncertainty shocks.

Google Trends data are freely available in real time. The first characteristic facilitates the replicability of scientific analysis, while the second one is consistent with the idea of constructing leading indicators, which is relevant for sharpening the identification of causal relationships. Ginsberg et al. (2009) use Google Trends data to predict influenza epidemics in the United States. Turning to economics, Choi and Varian (2012) exploit relevant search terms to predict car sales, unemployment claims, travel destination planning and consumer confidence. Baker and Fradkin (2017) use Google Trends data to compute a measure of job search to investigate the relation between unemployment insurance and job search in the United States. D’Amuri and Marcucci (2017) employ Google job searches and show that the Google Trends-based indicator they construct is the best leading indicator for the U.S. unemployment rate. Our paper focuses on the construction of uncertainty indices with Google Trends data.

While writing this paper, we became aware of a contribution by Bontempi et al. (2016). They construct an index of economic uncertainty for the United States using Google Trends data, contrast it with alternative measures of uncertainty, and use their measure in a VAR context to analyze the contribution of uncertainty shocks for the dynamics of employment and industrial production. With respect to them, we develop an uncertainty index also for Australia, focus on the unemployment rate as business cycle indicator given its central role for policymakers, and contrasts the role played by uncertainty shocks with that played by monetary policy shocks, which have been shown to have an influence on proxies of uncertainty (Pellegrino, 2017). Our contribution is also close to Baker et al. (2016), who construct an index of economic policy uncertainty (EPU) for a number of countries including the United States and Australia by searching uncertainty-related keywords conditional on a set of widely read country-specific newspapers, and to Moore (2017), who constructs an index of economic uncertainty for Australia based on keywords searches conditional on a set of Australian newspapers. We construct novel indices of uncertainty with a similar keyword-related search strategy but conditional on the freely available Google Trends database.

The structure of the paper is the following. Section 2 offers a brief presentation of our GTU index. Section 3 documents our VAR-related findings. Section 4 concludes.

2. GTU index

Construction of the GTU index. The construction of Google Trends uncertainty indices is based on the assumption that economic agents, represented by Internet users, look for online information when they are uncertain. This assumption implies that the search frequency of terms that may be associated to future, uncertain events is high when the level of uncertainty is high. To construct our country-specific indices, we proceed as follows. First, we subjectively select a broad set of keywords that are often cited in the Federal Reserve’s Beige Book for the U.S. and the Reserve Bank’s Statement on Monetary Policy in correspondence of uncertainties about future economic conditions. Examples of these words are “bankruptcy”, “stock market”, “economic reforms”, “debt stabilization”. Conditional on a given geographical area, which is, the United States for the U.S. GTU index and Australia for the Australian index, we then use Google Trends to recover the search frequency of each of these words in the sample of interest. Then, we aggregate the outcome of all individual searches per each month at a country level, therefore obtaining our GTU indices. Our Appendix A offers a detailed explanation on the construction of our indices. As anticipated in the Introduction, given the nature of the search term, the GTU indices are interpreted as measures of business uncertainty.

Fig. 1 shows the evolution of our indices for the United States and Australia. The major spikes in the U.S. index are associated with terms like “United States Congress”, “bankruptcy”, “White House”, “stock market”, “health care reform”, “debt ceiling”. Unsurprisingly, these terms are related to policy decisions or events affecting financial markets that are likely to inject uncertainty in the economic system and, therefore, suggest a pause in investment and a fall in labor demand by entrepreneurs. Similarly, spikes in the index for Australia are associated to words like “exchange rate”, “United States dollar”, “Euro”, “loan”, and “Australian Security Exchange”, which are naturally connected with entrepreneurs’ decisions.

Correlation with existing measures of uncertainty. Table 1 shows the correlation between the U.S. GTU index and a variety of different proxies for uncertainty proposed in the literature and available at monthly frequency. We consider the VXO used by Bloom (2009); the EPU index constructed by Baker et al. (2016); the macroeconomic uncertainty index proposed by Jurado et al. (2015); the financial uncertainty index constructed by Ludvigson et al. (2016); the subjective interest rate uncertainty proposed by Istrifi and Mouabbi (2017); the categorical measure of monetary policy-related uncertainty produced by Baker et al. (2016); the real-time, real activity related uncertainty index constructed by Scotti (2016); and the real-time measure of uncertainty based on the distribution of the forecast errors of real GDP constructed by Rossi and Sekhposyan (2015). These correlations are all positive and range from 0.04 (with Rossi and Sekhposyan’s index) to 0.63 (with Jurado et al.’s). The low correlation between our index and Rossi and Sekhposyan’s may be explained by the different frequency at which these indicators are constructed, as well as the fact

---

1 The dating of the business cycle for a number of industrialized countries is available at the Economic Cycle Research Institute’s webpage: https://www.businesscycle.com/.

2 Moore (2017) combines the information related to newspapers with that coming from stock market volatility, analyst earnings forecast uncertainty, and GDP growth forecast dispersion.

3 Our GTU indices are seasonally adjusted via the X-13ARIMA-SEATS Seasonal Adjustment Program downloadable from the U.S. Census Bureau’s website.

4 Our methodology allows us to search for the words whose search frequency is the highest per each given spike of our indices. The caption of Fig. 1 reports the three words with the highest relative search frequency per each given spike.

5 Scotti (2016) index is available at daily frequencies. We transform it to a monthly index by taking within-month averages. Rossi and Sekhposyan (2015) index is available at quarterly frequencies. We then correlate such measure with quarterly counterparts of the other proxies for uncertainty cited in the text by taking within-quarter averages of monthly values. We consider the one-year ahead version of Rossi and Sekhposyan’s index, which leads to higher correlations with the other proxies considered here than its nowcast counterpart (full set of results available upon request).
that our index likely captures information over and above the one related to the forecast of real GDP per se. The much higher correlation (0.49) with Scotti (2016) real-time index, which is constructed by exploiting a broad set of real activity indicators, corroborates this statement. As regards Australia, fewer proxies for uncertainty are available. The correlation between our GTU index and Baker et al.’s is 0.50, that of Moore’s index is 0.60, while that with the stock market volatility is 0.59. Overall, we interpret the positive sign of these correlations as reassuring as far as the quality of our uncertainty indices is concerned. Looking at other indices, it is of interest to notice that the Baker et al. (2016) overall EPU index is negatively correlated (−0.22) with the Istrefi and Mouabbi (2017) one, but its monetary-policy specific categorical version is actually positively correlated (0.28) with Istrefi and Mouabbi’s. This finding speaks in favor of the ability of Istrefi and Mouabbi’s index to capture the monetary policy-specific uncertainty dimension.

3. VAR Evidence

**VAR Investigation.** We identify the macroeconomic effects of GTU shocks by modeling selected macroeconomic series with country-specific VAR models which read: $\mathbf{X}_t = \Pi(\mathbf{L}) \mathbf{X}_{-t} + \mathbf{r}_t$, where $\mathbf{X}$ is a set of endogenous variables, $\Pi$ is a matrix of VAR coefficients capturing the dynamics of the system, and $\mathbf{r}_t \sim N(0, \Omega)$ is the vector of reduced-form residuals having zero-mean and variance-covariance matrix $\Omega$. $\Omega$ is estimated via OLS. To make sure that GTU shocks are orthogonal to the other stochastic elements in the econometric framework, we model the impulse vector responsible of the on-impact response of the variables in the vector $\mathbf{X}_t$ by employing a Cholesky-decomposition of the reduced-form variance-covariance matrix $\Omega$. The vector of U.S. data is $\mathbf{X}^{US}_t = \{GTU^{US}_t, u^{US}_t, p^{US}_t, SR^{US}_t\}$, where $GTU_t$ is our GTU index, $u_t$ stands for the civilian unemployment rate, $p_t$ stands for the price index (modeled in log-level and multiplied by 100), and $SR_t$ stands for the shadow-rate constructed by Wu and Xia (2016), which accounts for unconventional policies during the zero lower bound (ZLB) period. As regards Australia, we model $\mathbf{X}^{AU}_t = \{GTU^{AU}_t, u^{AU}_t, p^{AU}_t, R^*_t, \epsilon^{AU}_t\}$, where the first three variables have the same interpretation given above, $R_t$ is the cash rate, and $\epsilon_t$.

Following Bloom (2009), we compute the within-month volatility of stock market returns and use it as a proxy for financial uncertainty. Source of the data: Datastream. We use this measure of volatility because it covers the whole 2004M1–2016M12 sample. The SP ASX200 VIX for Australia is available starting from January 2008. The correlation between our measure of stock market volatility and the VIX in the sample 2008M1–2016M12 is 0.82.

### Table 1: Measures of uncertainty: Correlation with GTU indices

<table>
<thead>
<tr>
<th>United States</th>
<th>GTU</th>
<th>BBD</th>
<th>VVO</th>
<th>JLN</th>
<th>LMN</th>
<th>IM</th>
<th>BBD MP</th>
<th>SC</th>
<th>RS</th>
</tr>
</thead>
<tbody>
<tr>
<td>BBD</td>
<td>0.28</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>VVO</td>
<td>0.54</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>JLN</td>
<td>0.63</td>
<td>0.54</td>
<td></td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>LMN</td>
<td>0.53</td>
<td>0.20</td>
<td>0.76</td>
<td></td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>IM</td>
<td>0.38</td>
<td>−0.22</td>
<td>0.41</td>
<td>0.62</td>
<td>0.31</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>BBD MP</td>
<td>0.37</td>
<td>0.57</td>
<td>0.47</td>
<td>0.26</td>
<td>0.35</td>
<td>0.28</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SC</td>
<td>0.40</td>
<td>0.17</td>
<td>0.55</td>
<td>064</td>
<td>0.48</td>
<td>0.43</td>
<td>0.16</td>
<td>0.08</td>
<td>1</td>
</tr>
<tr>
<td>RS</td>
<td>0.04</td>
<td>0.16</td>
<td>0.38</td>
<td>0.45</td>
<td>0.48</td>
<td>0.43</td>
<td>0.16</td>
<td>0.08</td>
<td>1</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Australia</th>
<th>GTU</th>
<th>BBD</th>
<th>Moore</th>
<th>VOL</th>
</tr>
</thead>
<tbody>
<tr>
<td>BBD</td>
<td>0.50</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Moore</td>
<td>0.60</td>
<td>0.82</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>VOL</td>
<td>0.59</td>
<td>0.46</td>
<td>0.70</td>
<td>1</td>
</tr>
</tbody>
</table>

**Fig. 1.** Google trends uncertainty (GTU) indices. Sample: 2004M1–2016M12. Indices constructed by searching queries related to a battery of country-specific keywords as explained in the text and normalized to have mean = 100 and standard deviation = 30. Keywords for each identified spike, United States: A = gas price, bankruptcy, United States Congress; B = United States Congress, Federal Deposit Insurance Corporation, bankruptcy, recession, Federal Reserve System; C = bankruptcy, United States Congress, stock market; D = reform, health care reform, United States Congress; E = debt ceiling, bankruptcy, National debt of the United States; F = fiscal cliff, United States Congress, bankruptcy; G = United States Congress, reform, stock market; H = minimum wage, stock market, bankruptcy; I = minimum wage, United States Congress, stock market. Keywords for each identified spike, Australia: A = Australian security Exchange, Centrelink, Exchange rate; B = Australian security Exchange, Centrelink, United States Dollar; C = Centrelink, Euro, United States Dollar; D = Loan, Centrelink, United States Dollar.
is the trade-weighted nominal exchange rate (modeled in log-level and multiplied by 100).\footnote{The CPI index for Australia is available at quarterly frequency. We constructed a monthly version of such index by interpolating the quarterly observations via cubic spline.} We use unemployment to capture the response of real activity due to its availability at monthly frequency (as opposed to, say, real GDP or investment, which are available at quarterly frequency). Apart from the GTU indices, which were constructed via the Google Trends function, all U.S. data were downloaded from the Federal Reserve Bank of St. Louis’ website. Australian data were downloaded from the Reserve Bank of Australia’s website, with the exception of unemployment, which was downloaded from the Australian Bureau of Statistics’ website.

We focus on the sample 2004M1–2016M12. The beginning of the sample refers to the first month of availability of the Google Trends data. The VARs feature equation-specific constants and linear trends and are estimated with three lags.

Fig. 2 shows the impulse responses to a one-standard deviation country-specific GTU uncertainty shock. Such a shock generates a significant recessionary and deflationary (temporary) reaction in the United States, with a maximum absolute increase in the unemployment rate equal to about 0.15% and a peak decrease of about 0.2% of the price level. This macroeconomic responses, which are in line with those documented by Leduc and Liu (2016) using alternative proxies for uncertainty, are associated to a temporary drop in the federal funds rate. A one-standard deviation GTU shock in Australia generates much more moderate responses of prices and unemployment, with a decrease in the price level equal to 0.02% and a peak variation in unemployment equal to 0.03%. The peak response of the policy rate – a drop in the cash rate of 8 basis point – is comparable to the one in the U.S., possibly also due to the temporary depreciation of the nominal exchange rate (−0.7%). This evidence is similar to Moore’s (2017), who finds moderate real effects of uncertainty shocks in Australia using his measure of economic uncertainty.

Fig. 2. Impulse Response Functions to a GTU Shock. Sample: 2004M1–2016M12. VAR(3) estimated with a constant and a linear trend. Blue solid lines and dashed black ones: Point estimates and 68% percent confidence bands.

Table 2 documents the 4-year ahead forecast error variance decomposition analysis focusing on the contribution of the uncertainty and monetary policy shocks. As regards the United States, the contribution of uncertainty shocks to the volatility of unemployment is as high as 18%, and it is much larger than that of monetary policy shocks (about 5%). Differently, Australia is characterized by a much milder contribution of uncertainty shocks to the volatility of unemployment (about 6%), and a substantial one by monetary policy shocks (almost 30%).

Wrapping up, our main result is that uncertainty shocks are found to play a much larger role as drivers of the business cycle than monetary policy shocks in the United States. Shocks to the shadow rate are in part associated to unconventional monetary policy actions. Such actions are captured by the difference between the shadow rate and the federal funds rate. This difference is non-zero during the 2009M1–2015M11 time-span. Hence, the comparison with the effects of Australian monetary policy shocks should be taken with a grain of salt. If anything, modeling unconventional monetary policy shocks – engineered to push the United States out of their period of stagnant growth – works in favor of downplaying the effect of other shocks – uncertainty shocks included – on unemployment in our VAR. In other words, our results in favor of the relatively larger real effects played by uncertainty shocks as opposed to monetary policy shocks in the United States would likely be even stronger if we conditioned on conventional monetary policy shocks only.

Table 2. Forecast error variance decomposition. 4 year-ahead forecast error variance decomposition. The figures reported in the table refer to the point estimates of the baseline models.

<table>
<thead>
<tr>
<th>United States</th>
<th>Shock/Variable</th>
<th>GTU</th>
<th>PT</th>
<th>UT</th>
<th>RT</th>
<th>ET</th>
</tr>
</thead>
<tbody>
<tr>
<td>ϵGTU</td>
<td>83.39</td>
<td>20.84</td>
<td>18.46</td>
<td>1.88</td>
<td>–</td>
<td></td>
</tr>
<tr>
<td>ϵP</td>
<td>2.97</td>
<td>12.49</td>
<td>4.70</td>
<td>26.53</td>
<td>–</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Australia</th>
<th>Shock/Variable</th>
<th>GTU</th>
<th>PT</th>
<th>UT</th>
<th>RT</th>
<th>ET</th>
</tr>
</thead>
<tbody>
<tr>
<td>ϵGTU</td>
<td>51.49</td>
<td>0.47</td>
<td>6.41</td>
<td>8.31</td>
<td>2.96</td>
<td></td>
</tr>
<tr>
<td>ϵP</td>
<td>23.65</td>
<td>2.17</td>
<td>29.50</td>
<td>49.06</td>
<td>7.86</td>
<td></td>
</tr>
</tbody>
</table>
in the U.S. than in Australia. We interpret this result in light of the literature cited in the Introduction, which points to uncertainty shocks as being particularly harmful for the business cycle when the economy is already weak and conventional monetary policy cannot operate because of the ZLB.

Robustness checks. Our results are robust to: (i) ordering uncertainty last in the vectors, which enables us to control for the possible role played by contemporaneous variables in the VAR in affecting uncertainty; (ii) modeling different VAR lags; (iii) controlling for global uncertainty pressures proxied by the Global Economic Policy Uncertainty index developed by Davis (2016); (iv) controlling for uncertainty in China (proxied by the EPU uncertainty index developed by Baker et al., 2016). The controls in exercises (iii) and (iv) are added to the baseline vectors and ordered first. All these checks are documented in our Appendix A.

4. Conclusions

This paper constructs novel measures of uncertainty based on Google Trends for the United States and Australia. Such measures are shown to be positively correlated with alternative ones. VAR investigations point to exogenous variations in these measures as being able to predict movements in unemployment in the United States and, to a much lesser extent, in Australia. Future research could exploit the flexibility of the Google Trends approach to build measures of uncertainty for different countries or at a state level. A contribution along the latter avenue is Tran (2017).

Appendix A. Supplementary data

Supplementary material related to this article can be found online at https://doi.org/10.1016/j.econlet.2017.09.032.

References

Leduc, S., Liu, Z., 2016. Uncertainty Shocks are aggregate demand Shocks. J. Monetary Econ. 82, 20–35.