Crawford School of Public Policy

CAMA

Centre for Applied Macroeconomic Analysis

Uncertainty in a Disaggregate Model: a Data Rich Approach using Google Search Queries

CAMA Working Paper 83/2019
November 2019

Kalvinder Shields
Department of Economics, University of Melbourne
Centre for Applied Macroeconomic Analysis, ANU

Trung Duc Tran
School of Economics, University of Sydney
Centre for Applied Macroeconomic Analysis, ANU

Abstract

This paper estimates the impact of uncertainty shocks in a disaggregate model featuring state-level unemployment and uncertainty, which is measured using Google search data. We show that the disaggregate model captures important spillover effects which a model using aggregate data would overlook resulting in significantly different peak responses and time dynamic effects. We find the effect of uncertainty shocks on state-level unemployment is recessionary and heterogeneous. The importance of national factors in propagating the effect of uncertainty is also heterogeneous across states, and overall less relevant than state-level factors. These heterogeneous effects are found to be related to state-specific industry compositions
and the fiscal position.

Keywords

Google Trends uncertainty, uncertainty shocks, regional effects

JEL Classification

C32, E32

Address for correspondence:

(E) cama.admin@anu.edu.au

ISSN 2206-0332

Uncertainty in a Disaggregate Model: a Data Rich Approach using Google Search Queries

Kalvinder Shields* and Trung Duc Tran†‡

November 2019

Abstract

This paper estimates the impact of uncertainty shocks in a disaggregate model featuring state-level unemployment and uncertainty, which is measured using Google search data. We show that the disaggregate model captures important spillover effects which a model using aggregate data would overlook resulting in significantly different peak responses and time dynamic effects. We find the effect of uncertainty shocks on state-level unemployment is recessionary and heterogeneous. The importance of national factors in propagating the effect of uncertainty is also heterogeneous across states, and overall less relevant than state-level factors. These heterogeneous effects are found to be related to state-specific industry compositions and the fiscal position.

JEL-codes: C32, E32

Keywords: Google Trends uncertainty, uncertainty shocks, regional effects

*Professor, Department of Economics, University of Melbourne. Email: k.shields@unimelb.edu.au

†Postdoctoral Research Fellow, School of Economics, University of Sydney. Email: trungduc.tran@sydney.edu.au

‡The authors would like to thank Giovanni Caggiano, Efrem Castelnuovo, Andrea Carriero, Punnoose Jacob, Mariano Kulish, Anella Munro, Viet H. Nguyen, Kian Ong, Giovanni Pellegrino, Guido Turnip, Tugrul Vehbi, Benjamin Wong; and participants at the University of Melbourne Brown Bag Seminar, the 2017 International Association of Applied Econometrics (IAAE) Conference in Sapporo, the 2017 Australian Conference of Economists in Sydney, the 2017 Workshop of the Australasian Macroeconomics Society in Canberra and the Reserve Bank of New Zealand Round Table for helpful feedback that has greatly improved this manuscript. The authors would also gratefully acknowledge financial support by the Australian Research Council via the Discovery Grant DP160102281, the University of Melbourne Faculty Grant & Research Scholarship and the IAAE Travel grant.

1 Introduction

Uncertainty is commonly identified as being a major driver in the slow recovery of real economic activity during and after recessions with uncertainty influencing the dynamics of economic activity. This can occur directly, through the interaction of uncertainty with attitudes towards risk and subsequent investment in the real and financial economy, or indirectly, through prolonging the effects of other recessionary shocks.[1][2] Whilst the focus of almost all of this literature has concerned the aggregate impact of uncertainty on real activity, the use of national measures of uncertainty however conceals the considerable heterogeneity across regional states due to variations in the composition of production, employment, demographics and the regulatory differences in the financial and public sector, for instance. Such variations prompt the need to consider the importance of measuring uncertainty at a more disaggregated regional level of an economy, the respective impact on economic activity at a more regional level, and accordingly, the implications for an aggregate view through a disaggregate lens accommodating potential interactions and spillovers that might mitigate or amplify the transmission of uncertainty across the economy.

This paper advances the uncertainty literature, therefore, in the following directions. First, the paper proposes a novel and credible measure of uncertainty for each of the 51 U.S. states through using Google Trends search data (GTU). The credibility of such an approach to measure uncertainty is illustrated in Castelnuovo and Tran (2017) and based in part due to their findings of substantial correlations of the GTU index for the United States as a whole with a variety of different proxies for uncertainty proposed in the literature and available at a monthly frequency.[3] However, the absence of an objective measure

[1]For discussion, see Bloom (2014), Fernández-Villaverde et al. (2011), or Ilut and Schneider (2014), for instance.

[2]The direct and indirect effect here may not mean the same as the exogenous or endogenous effect of uncertainty shocks.

[3]For instance, Castelnuovo and Tran (2017) consider the VXO used in 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. (2018); the subjective interest rate uncertainty proposed by Istrefi 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). With the exception of Rossi and

of uncertainty has meant that previous empirical studies conducted at the U.S. state-level have relied on observable proxy measures based on counting uncertainty-related words in newspapers and tax codes as in Shoag and Veuger (2016) or model-based econometrics estimates as in Mumtaz et al. (2018) and Mumtaz (2018). The use of Google search data provides real time, easily accessible and timely data to construct measures of uncertainty at the state level for the US economy. Second, this is the first paper to explicitly accommodate the dynamic interactions between uncertainty and economic activity across states and across time through a comprehensive and integrated framework that jointly models monthly uncertainty and the unemployment rate for 51 states whilst allowing for feedback between these variables.[4] The framework of analysis is the Global Vector Autoregressive modelling framework (GVAR), elaborated in, inter alia, Garratt et al. (2006), Dees et al. (2007) and Chudik and Pesaran (2016), and is able to complement the existing state-level studies as in Mumtaz et al. (2018), Mumtaz (2018) and Shoag and Veuger (2016) by (i) providing an extremely flexible method for characterising the evolution over time of the respective state-specific uncertainty and economic activity variables, and (ii) allowing for relatively complicated forms of interactions and spillover effects between these state-specific variables. Third, our proposed modelling framework allows us to provide original insights on the relative importance of national factors versus state-specific factors in their propagation of uncertainty shocks on unemployment dynamics at the state-level and their respective associations with a number of state-specific characteristics.

The construction of the uncertainty index is based on the premise that when individuals are uncertain about the future, the Internet is a likely source of reference for information; when there is uncertainty about particular topics, search frequencies for words related to such issues of interest are expected to be higher. This paper proposes

Sekhposyan (2015), all the correlations are high - and the low correlation with Rossi and Sekhposyan's measure is explained by the different frequency at which these indicators are constructed, and that the GTU index captures information over and above related to the forecast of real GDP per se. The much higher correlation (0.49) with Scotti's (2016) real-time index, which is constructed by exploiting a broad set of real activity indicators, corroborates this statement.

[4]Previous papers by Mumtaz et al. (2018) abstract from measuring uncertainty at the state-level choosing to focus only on the impact of national uncertainty on US states, whilst Mumtaz (2018) and Shoag and Veuger (2016) abstract from state-level feedbacks as well as time dynamics in the propagation of state-level uncertainty.

that a reliable proxy of uncertainty can be built upon the uncertainty-related keywords as mentioned in the 'Beige Books' as published by the Federal Reserve. The Beige Books gather information on current economic conditions based on interviews with key business contacts, economists, and market experts. Therefore, the GTU is likely to capture uncertainty relevant to the business environment. In addition, the GTU index constructed for the US as a whole is shown to be consistent with other existing uncertainty indices at the aggregate level (Bachmann et al., 2013; Baker et al., 2016; Caggiano et al., 2014; Istrefi and Mouabbi, 2017; Jurado et al., 2015; Ludvigson et al., 2018; Scotti, 2016; Shields et al., 2005). GTU indices constructed at the state level through the Google Search Tool allowing the isolation of searches for each state have credibility therefore.

The importance of disaggregation is relevant to consider in any modelling framework given that the basis of how the economy operates generally involves (highly) interconnected disaggregated units. However, there will be trade-offs concerning the level of disaggregation since a model at a high level of disaggregation may contain detail which may obscure the primary features of the model as well as contain errors which, in the aggregate, cancel out. Further trade-offs between models developed at different levels of disaggregation relate to the costs of collection, the accuracy of representation and their ability to adequately represent the relevant historical data. There is also, typically, little guidance on how to discriminate between such models. In this instance, disaggregate data using Google Search data is easily accessible as already mentioned. However, the question of whether a more disaggregate model adds significant information over an aggregate model will need some formal analysis. In this paper we base the choice of model depending on each of their abilities to predict aggregate variables of interest (see Pesaran et al. (1989) (PPK), Pesaran et al. (1994) (PPL) and Van Garderen et al. (2000) (GLP), for instance). PPK and PPL propose selection criteria allowing one to judge the statistical adequacy of (linear) aggregate and disaggregate specifications based on the prediction of aggregate series of interest.[5] Based on these selection criteria and using parametric bootstrapping methods to account for sample variation (see GLP), there is strong evidence

[5]GLP extend this work by considering the model selection problem when the underlying micro behavioural relations involve non-linearities as well as dynamics.

to support the usefulness of disaggregated data in models of uncertainty and economic activity for the US. The use of a misspecified aggregate model would therefore overlook relevant feedbacks and interactions in the analysis on the impact of uncertainty on the economy.

We find that the impact of aggregate uncertainty shocks on unemployment in the disaggregate model is lower compared to the aggregate model and the disemmination of the uncertainty shocks taking almost twice as long relative to the aggregate model. This suggests that the process of aggregation, in this case, omits essential dynamics concerning the propagation of uncertainty across states and thereby amplifies the persistence of these series whilst underestimating the propagation time of uncertainty. A one-standard-deviation aggregate uncertainty shock in the disaggregate framework causes aggregate unemployment rate to go up by 0.1 percentage points at peak. Such a shock in the aggregate model, on the other hand, causes the aggregate unemployment rate to go up by approximately 0.28 percentage points at peak. Although the peak effect is different from a statistical perspective, we still find an important effect of uncertainty shocks in both the models. In terms of the propagation of uncertainty, we find that the uncertainty shock takes approximately two years to work its way through a (misspecified) aggregate model and half the time compared to the disaggregate model which allows for statistically relevant feedbacks and interactions between states.

In addition, we find that the responses of unemployment in US states to an aggregate uncertainty shock are heterogeneous. The paper provides a narrative to this heterogeneity using various variables capturing state-specific characteristics such as industry composition, fiscal constraints, labour market and financial frictions in a post regression analysis. We find that the unemployment rates in states with a larger concentration of the manufacturing industries are affected more by uncertainty shocks, whilst a larger share of the mining industry mitigates the impact of uncertainty.

The use of the GVAR framework allows us to evaluate the relative importance of a national factor (i.e. the 'global' factor in the GVAR model) in driving economic activity at the state level (Garratt et al., 2018). On average, we find the national factors to be less important than the state-specific factors in their respective role in propagating the effects of uncertainty shocks on state-level unemployment, although there is significant heterogeneity across states. In general and, as anticipated, the presence of national factors serves to amplify the effects of uncertainty shocks on state-level unemployment. We then investigate the explanatory power of state-specific characteristics in the heterogeneous importance of the national factors. The results show the relative importance of national influences is greater in states where the real estate sector is greater - and reflects an integrated financial market for instance. In contrast, we find states with a more active fiscal policy experience a smaller influence of national factors in the impact and propagation of uncertainty shocks. This is illustrated by a negative and statistically significant relationship between the importance of national influences in propagating uncertainty shocks and the share of the government sector, the size of state-level debt, and the size of state-level spending.

The rest of the paper is organised as follows. Section 2 provides a brief overview of popular measures of uncertainty as well as findings of studies investigating the impact of uncertainty on the economy. Section 3 explains the construction of the uncertainty index. Section 4 presents the GVAR framework on which this analysis is based. Section 5 outlines the statistical methods for assessing the adequacy of the disaggregate model relative to its aggregate counterpart, investigates the dynamics of the GVAR model and estimates the impact of uncertainty of uncertainty shocks on economic activity. This section further explores the relative importance of national factors versus state-specific factors in propagating the effect of uncertainty shocks for each state through a decomposition analysis based on Garratt et al. (2018). Section 6 provides a subsequent regression analysis using state-specific characteristics to provide insights into the heterogeneous effects of uncertainty. Section 7 provides some concluding comments.

2 Related Literature

This paper relates to at least three strands of literature. First, there is a literature focusing on measuring uncertainty. In general, there are three approaches by which uncertainty could be proxied as described in Bloom (2014): (i) based on financial data as in Bloom (2009) or Caggiano et al. (2014); (ii) based on forecast and estimation as in Angelini et al. (2018), Bachmann et al. (2013), Carriero et al. (2018), Jurado et al. (2015), Istrefi and Mouabbi (2017), Leduc and Liu (2016), Ludvigson et al. (2018), Shields et al. (2005), Scotti (2016) and Rossi and Sekhposyan (2015) ; and (iii) based on news as in Baker et al. (2016) or online search data as in Bontempi et al. (2016) and Castelnuovo and Tran (2017).

The first approach often makes use of high frequency and observable financial data such as the VIX. The drawback however is such an index only considers financial uncertainty. The estimation-based uncertainty index is statistically sound. However, the estimated measure of uncertainty is constrained by the structured form of the respective econometric frameworks and typically rely on official economic indicators that are released with lags and which may require subsequent revisions. Another method in the second approach is through the use of business or consumer confidence surveys. For example, Leduc and Liu (2016) construct an uncertainty index from the Michigan Survey of Consumers. Similarly, Bachmann et al. (2013) obtain uncertainty from business outlook data in the manufacturing industry in Germany and in the U.S. Despite the sound economic theory behind this approach, uncertainty proxied by disagreement amongst professional forecasters could just be a measure of divergence in opinion among forecasters rather than the underlying level of uncertainty in the economy.

Finally, being the most notable proxy in the third approach, Baker et al. (2016) construct their Economic Policy Uncertainty index (EPU) by counting the frequency of certain uncertainty-related keywords appearing in newspapers; this is directly observable and can relate to specific events that cause uncertainty to rise. The shortcoming of the index constructed by counting the frequency of newspaper articles mentioning economic uncertainty tends to be reflections of journalists' thoughts on the economy and it is not necessarily a representation of uncertainty in the whole economy. Lastly, and similar to this approach, Bontempi et al. (2016) also use online search data to construct their uncertainty index and focus on evaluating the online search uncertainty versus other existing indices. It is this vein of the literature that we focus on and contribute to Bontempi et al. (2016), for instance, by considering the construction and evaluation of disaggregated uncertainty indices within a disaggregate framework of the economy.

The second strand of literature estimates the impact of uncertainty. The literature can be traced back to the theoretical model by Bernanke (1983) who links high uncertainty and the incentive to delay investment. The recent Great Recession has renewed the interest in uncertainty, led by the seminal work by Bloom (2009). In short, the theoretical models in this literature are based on adjustment frictions. The interaction between high uncertainty and non-convex adjustment frictions may cause firms to behave more cautiously; thus firms "wait and see" before making any investment or hiring decision. The empirical literature provides strong evidence that the effect of uncertainty shocks are negative (see Bachmann et al. (2013), Jurado et al. (2015) and Leduc and Liu (2016)).

The estimates reported in those papers typically focus on the impact on the U.S national data. Less is known about the state-level effects of uncertainty. Shoag and Veuger (2016) use a similar approach to Baker et al. (2016) to construct state-level uncertainty indices in the U.S, and find a causal role for uncertainty in increasing unemployment at the state level over the 2007 – 2009 recession period, working within a cross-sectional regression framework. Mumtaz et al. (2018), motivated by the structural differences in each state, find that the impact of national uncertainty on state-level real income differs across states. Mumtaz (2018) finds that that increases in uncertainty do have an adverse impact on real income, employment and unemployment by using variation in the effect of US-wide uncertainty on state-level uncertainty to identify the impact of this shock on real economic activity.

To the best of the authors' knowledge, this study along with Bontempi et al. (2016) are the first to use Google search data to construct an uncertainty index. [6] The online search data literature mostly involves forecasting, where the aim is to maximise the correlation between a chosen set of search term to the variable of interest. Ginsberg et al. (2009), for instance, were the first to use Google search data to forecast influenza outbreaks.

[6] Castelnuovo and Tran (2017) on the other hand use Google search data to measure aggregate uncertainty in the US and Australia.

The authors find a positive correlation between search queries and doctor visits, thereby developing an early influenza outbreak warning system. In the economics literature, early work by Choi and Varian (2009) finds that relevant search terms can be used to predict car sales, unemployment and many other economic variables in the US. D'Amuri and Marcucci (2017) also suggest that the Google job search index is the best leading indicator for the U.S unemployment rate.

Lastly, the final strand of literature relates to the GVAR model in the spirit of Dees et al. (2007). Such a framework is suitable for the purpose of investigating the state-level effect of uncertainty because it allows for linkages and interdependencies between the state-level factors and the national common factors. Whilst there is no established literature studying the dynamic state-level impact of uncertainty on economic activity, explicitly allowing for interdependencies and feedbacks between state-level uncertainty and economic activity, more generally or within the GVAR modelling framework, Cesa-Bianchi et al. (2014) and Garratt et al. (2018) model global uncertainty using the GVAR model. The only known application of the GVAR model at the U.S state level is by Vansteenkiste (2007) who studies the effect of house price spillovers across U.S states.

3 Measuring Economic Uncertainty

3.1 Background

The construction of the uncertainty index is based on the premise that when economic agents, represented by internet users, are uncertain about the future, they tend to look for information on the internet. Under this assumption, the search frequency would be high when the level of uncertainty for a certain topic is high. Google search data is well-suited to be the measure of economic uncertainty due to its representative power. According to comScore, Google has been dominating the online search market in the U.S, where its market share has risen from 56 % in 2004 to 65 % in 2016. By exploiting this data rich environment, the uncertainty index is able to capture the level of uncertainty represented by searchers who are potentially concerned or affected by the state of the economy. The list of search terms are abstracted from the Federal Reserve Beige Book. Due to the nature of the search terms, this uncertainty index cannot, therefore, be interpreted as a general index of economic uncertainty but rather be a measure of uncertainty for business owners and professionals. It reflects the concern of business owners and professionals on economic policy, economic conditions, as well as unexpected economic uncertainty from non-economic activity.

3.2 Construction of the GTU Index

This section provides a general representation of the construction of the index. An intuitive example of how to construct a simplified uncertainty index can be found in the Appendix. Data from Google Trends gives the frequency of a particular search term relative to the total search volume ranging between 0 and 100, rather than the absolute search frequency – which is not made available for privacy purposes. To construct this frequency, each raw data point, $R_{\omega,t,i}$, where ω denotes the search term, at time t, for the location it represents, i, is divided by the total searches, $\overline{S}$: $S_{k,t,i} = \frac{R_{\omega,t,i}}{\overline{S}_{t,i}}$.[7] For exposition purposes, the subscript i is dropped to simplify notation for now.[8] The resulting numbers are then scaled to a range between 0 and 100, where 100 represents the point where the search frequency is highest, $FI_{\omega,t} = \frac{100}{M} S_{\omega,t}$ and $M = \max\{S_{\omega,1}, S_{\omega,2}, \ldots, S_{\omega,\overline{S}}\}$. Data is excluded if searches are made by very few people. The downside of Google search data, as noted by Choi and Varian (2009) is that the exact replication of the data is not feasible due to sampling variability, especially for small volume search terms. This, however, is not an issue in this study since small volume search terms play minimal role after the aggregation process which is outlined in what follows.

The chosen method of aggregating aims to reflect the true search volume relating to each respective search term and, accordingly, a term which is searched more frequently will hold a larger weight in the final aggregated index. As Google only allows the input of a maximum of five different search terms in the Google Trends search engine at any one

[7]Depending on specific search term, the frequency could either be daily, weekly or monthly. This paper deals with the monthly search frequency.

[8]Subsequently, for state-level uncertainty, i, will denote the name of the respective state.

time, a benchmark term is chosen for the purpose of aggregation since the search frequency provided by Google Trends, $FI_{\omega,t}$ will alter depending on the choice of the other search terms included in that particular search. This is implemented in the following manner. First, a benchmark term is chosen and entered into the Google Trends search engine together with another search term. The search frequency for the benchmark term in this search is $FI^1_{\text{benchmark},t}$ and the output for the other term is $FI^1_{\omega,t}$. Second, this process is repeated for the rest of the list of search terms. The frequency of the chosen benchmark term when entered into Google Trends with another search term, $FI^2_{\text{benchmark},t}$, will not be necessarily the same as the benchmark value in the first step, $FI^1_{\text{benchmark},t}$ as the highest search term in the new combination is automatically set to take a maximum of 100. The search frequency for the remaining search terms in Step 2 is $FI^2_{\omega,t}$. The true frequency of new search terms (conditional on the search frequency of the benchmark term in Step 1), $FI^1_{\omega,t}$, is calculated by using the adjusting ratio, $FI^1_{\omega,t} = FI^2_{\omega,t} \times \frac{FI^1_{\text{benchmark },t}}{FI^2_{\text{benchmark},t}}$.[9]

The Google Trends uncertainty index, GTU_t is the sum of all the scaled search terms where $\overline{\omega}$ is the total number of search terms:[10]

$$GTU_t = \sum_{\omega=1}^{\overline{\omega}} FI^1_{\omega,t} \tag{1}$$

The construction of the index involves identifying which search terms are related to the level of uncertainty in the economy. This is similar to the approach undertaken in Baker et al. (2016) which involved counting the frequency of uncertainty-related words in newspapers. The list of search terms in this paper makes use of the Federal Reserve Bank's Beige Books. The Beige Book is a summary of economic conditions in the U.S prepared by the 12 Federal Reserve Districts. Each Federal Reserve Bank gathers anecdotal information on current economic conditions in its District through reports from Bank and Branch directors and interviews with key business contacts, economists, market experts,

[9]Bontempi et al. (2016) use both Bayesian aggregation and the common-term aggregation approach to find similar correlation between the two approaches.

[10]The advantage of this approach is the final list of search term is likely to be exhaustive since some terms are searched more; thus making other terms less important. However, it also raises a potential concern relating to the true importance of those dominating search terms in determining the level of uncertainty in the economy. The appendix looks at this issue and finds this is not a concern when using this approach.

and other sources. Therefore, the Beige Book contains valuable information on the factors which contribute to economic agents' perceptions of uncertainty. When uncertainty is mentioned in a passage therefore, the passage is examined in more detail to determine the factors which are causing the respective economic agents being interviewed to reveal their uncertainties. Examination of the Beige Book over the past 20 years gives sufficient information on the factors which are accordingly being associated with uncertainty.

Table 1 presents the full list of search terms used in this paper. Given these search terms, it is reasonable to appreciate that these search queries are arising from managers, economists and other professionals rather than from the general public. Consequently, uncertainties are typically related to the banking and finance sector, economic and business conditions, price levels, job market conditions, fiscal and monetary policy, change in regulations and the housing market.[11]

Figure 1 plots the time-varying uncertainty index (GTU_{it}) from 2006M01 - 2018M03, for selected states (i), along with the national uncertainty index, where the national uncertainty index is constructed as the population-weighted average of 51 individual state-level uncertainties in the US, constructed according to the approach described above. The list of all U.S. states are given in Table 2. Uncertainty at state level and at the national level peaks during the Great Recession and returns to more stable levels post-recession. Both state-level uncertainties and the national level uncertainty show considerable co-movements for much of the time although there are period in which there is considerable divergence - in the cases of California and New York, for example. The deviation in New York in June 2011, for instance, is caused by search terms involving austerity, debt ceiling, health care reform and unemployment benefits. The list of keywords associated with each period of a major deviation can be found in the Appendix.

Conceptually speaking, the idiosyncratic characteristics of uncertainty measures across different states are driven by differences in perceptions of uncertainty. However, it is also the case that there are two factors that can also contribute to such idiosyncratic behaviour.

[11]These categories are more or less similar to Baker et al. (2016) who have the following policy categories: Fiscal policy, monetary policy, health care, national security and war, regulation, foreign sovereign debt and currency crisis, entitlement programs and trade policy.

First, the share of internet users are not homogeneous across states, for instance, 65% of the population use the internet in Alabama, while this statistics reaches 80% in California. Second, the share of occupations across states may also be a major factor. Since the list of search terms used in this paper is more associated with the business environment and its respective economic decision-makers, in the states where there is a higher share of business activities, for instance, individuals in these states would relatively be more concerned about these keywords than those in states with a lower share of business-related activities.

4 The Modelling Framework

The GVAR framework provides an effective way of modelling interactions and feedbacks at a disaggregate state level; such a framework is not only able to interlink the U.S economy via a common uncertainty channel effect but also takes into account the unobserved interactions. The GVAR model is implemented in two steps. In the first step, state-specific models are estimated conditional on the rest of other states. These individual state-specific models are estimated in which the state-specific variables are related to the associated weighted-averages of other states' variables. In the second step, these individual state models are then linked and solved simultaneously as one large global VAR model (GVAR).

Assume there are N states, indexed by $i = 1 \ldots, N$. Each state i features k_i variables observed during the investigation period $t = 1, 2 \ldots T$. Let x_{it} denote a $k_i \times 1$ vector of variables specific to cross-section unit i in period t, and let x_t denote a $k \times 1$ vector of all variables in the model. Each individual model explains the specific state variable, x_{it}, conditional on the cross-section averages of all the other state variables, denoted, x^*_{it}, collected in the $k^*_i \times 1$ vector. In essence, x_t is the vector of the local state i variables and x^*_t denotes the vector of the variables capturing external influences outside of state i; these are commonly referred to as the ' *starred*' variables.

Each state is modelled as a VARX* model, where it is assumed that the endogenous state variables x_{it} are related to x^*_{it}, where the '*starred*' variables enter each individual model exogenously. The endogenous variables in most state-specific models include state-specific uncertainty, us_{it}, which are constructed using the methodology outlined in Section 3 and state-specific unemployment denoted as ur_{it}.

The '*starred*' variables are constructed as:

$$us^*_{it} = \sum_{j=0}^{N} w_{i,j} us_{jt}, \quad ur^*_{it} = \sum_{j=0}^{N} w_{i,j} ur_{jt}$$

for state i, where the weights $w_{i,j}$, $j = 0, 1, \ldots N$ captures the importance of state i to state j and $w_{i,j} = 0$ for $i = j$. Population weights are used to construct the link matrix.[12]

For each state $i \in 1, \ldots, N$, a VARX* model is estimated as:

$$x_{it} = \alpha_{i0} + \sum_{l=1}^{p_i} \Phi_{il} x_{i,t-l} + \sum_{l=0}^{q_i} \Lambda_{il} x^*_{i,t-l} + u_{it} \tag{2}$$

where Φ_{il} are $k_i \times k_i$ coefficient matrices relating to own respective states, Λ_{il} are $k_i \times k^*_i$ matrices for 'foreign' states, and p_i and q_i are the number of lags for each of these, respectively, and Σ_u denotes the variance-covariance matrix of u_{it}.

Although the estimation is carried out for each individual model, the GVAR model needs to be solved simultaneously for all endogenous variables globally. Once each state i model is estimated, the individual VARX* model for $i = 1, \ldots, N$ can be stacked together and solved as one global system and written as:

$$A_{i0} z_{it} = a_{i0} + \sum_{l=1}^{r_i} A_{il} z_{it-l} + u_{it} \tag{3}$$

where: $z_{it} = (x_{it}, x^*_{it})$, $r_i = \max\{p_i, q_i\}$, $A_{i0} = (I_{ki}, -\Lambda_{i0})$ and $A_{il} = (\Phi_{il}, \Lambda_{il})$ for $l =$

[12]The GVAR model is able to link the individual model together as a global system by using some intuitive weighting matrices that determine the interconnectedness among the individual models such as GDP weights, population weights or even distance weights. In this sense, the aggregate variable in the GVAR model is constructed from individual variables rather than being a true aggregate variable. Population weights are chosen because the correlation between the true US unemployment rate and the aggregated unemployment rate, and the correlation between the true US uncertainty and the aggregated uncertainty are higher relative to when other weighting matrices are used. The results are robust to when we use GDP weights, equal weights or distance weights. See the Appendix for further detail.

$1, \ldots, r_i$. The key to solving the model by noting the following identity:

$$z_{it} = W_i x_t \tag{4}$$

where $x_t = (x_{0t}, x_{1t}, \ldots, x_{Nt})'$ is the $k \times 1$ vector that collects all the endogenous variables of the system, and W_i is a $(k_i + k_i^*) \times k$ matrix that are defined by the weighting matrix w_{ij}. It follows that:

$$A_{i0} W_i x_t = a_{i0} + \sum_{l=1}^{r_i} A_{i,l} W_i x_{t-l} + u_{it} \text{ for } i = 0, 1, \ldots, N \tag{5}$$

and these individual models are then stacked to yield the model for x_t given by:

$$G_0 x_t = \alpha_0 + \sum_{l=1}^{r} G_l x_{t-l} + u_t \tag{6}$$

where

$$a_0 = \begin{bmatrix} a_{1,0} \\ a_{20} \\ \vdots \\ a_{N,0} \end{bmatrix}, G_0 = \begin{bmatrix} A_{1,0} W_1 \\ A_{2,0} W_2 \\ \vdots \\ A_{N,0} W_N \end{bmatrix}, G_l = \begin{bmatrix} A_{1,l} W_1 \\ A_{2,l} W_2 \\ \vdots \\ A_{N,l} W_N \end{bmatrix}, u_t = \begin{bmatrix} u_{1,0} \\ u_{20} \\ \vdots \\ u_{N,0} \end{bmatrix} \text{for } j = 1, \ldots, r$$

and $r = \max\{r_i\}$ for $i = 1, \ldots, N$. Since G_0 is a known non-singular matrix, premultiplying by G_0^{-1}, the GVAR(r) model is obtained as:

$$x_t = b_0 + \sum_{l=1}^{r} F_l x_{t-1} + \epsilon_t \tag{7}$$

where

$$b_0 = G_0^{-1} a_0, \;\; F_j = G_0^{-1} G_j \;\; \epsilon_t = G_0^{-1} u_t \text{ for } j = 1, \ldots, r$$

and the covariance matrix is given by $\Sigma_\epsilon = E(\epsilon_t' \epsilon_t)$. The GVAR framework therefore allows interactions among different states through three channels; (i) the contemporaneous effect of x_{it}^* on x_{it} and its lagged values; (ii) the effect of the common exogenous variable on the state-specific variable; and (iii) the contemporaneous dependence of shocks in state i on the shocks in state j.

5 Modelling Disaggregated Unemployment and Uncertainty Fluctuations in the U.S.

5.1 Data

This GVAR framework captures the spillover and feedback effects between uncertainty and unemployment within and across states. The model is estimated using two variables for each state: state-level uncertainty and the state-level unemployment for 51 U.S. states from 2006M01 to 2018M03. Uncertainty for each state is measured according to the method outlined in Section 3 and is in a levels form. Unemployment is measured as the quarterly change in the unemployment rate and is sourced from the Federal Reserve Bank of St Louis FRED Database.

Figure 2 shows the cross-state correlation plot for uncertainty (top panel) and the quarterly change in the unemployment rate (bottom panel). The Figure shows a considerable degree of comovement between uncertainty across states and unemployment across states. There are some exceptions, however; uncertainty in Delaware, (Washington) D.C, Vermont and Virginia do not seem to follow the comovement for the rest of the economy. Regarding unemployment, only unemployment in Louisiana has a minor correlation with unemployment in the other states and this seems mainly due to high unemployment experienced in 2006 caused by Hurricane Katrina.[13]

Due to the short time span of the Google search data, the maximum p_i is 2, and the maximum q_i is 1. Based on the Akaike Information Criterion, a VARX*(1,1) is fitted to most states.[14] A specification search is also performed on the coefficients and insignificant

[13]In the Appendix, we find the main results are robust to the removal of these states.

[14]The typical approach in the literature relies is based on the vector autoregressive error correction model (VECM) which allows for cointegration by assuming the *starred* variables to be weakly exogenous I(1) as in Dees et al. (2007). However, in this case there is no clear prior knowledge that uncertainty and unemployment would necessarily be cointegrated in the long run. The individual model is therefore estimated as a flexible VAR model.

coefficients for which the absolute value of the t-ratio being less than one are excluded from the model.[15]

5.2 Assessing the Adequacy of the Disaggregate Model

5.2.1 The Joint Significance of the Foreign Variables

GVAR models allow for potentially important cross-state interaction information that cannot be captured in more aggregate models. As described earlier, the presence of the *starred* or foreign variables allow uncertainty and unemployment to be inter-connected across states that are represented by the coefficient matrices Λ. To investigate the statistical importance of these variables however, the joint significance of cross-state interactions is calculated through an F-test of the joint significance of *starred* uncertainty variables and *starred* unemployment variables in each of the uncertainty and the unemployment equations for each of the individual state models.

Table 3 reports the F-test statistics associated with testing of the following null: H_0 : $\Lambda_{i0}^{\tau} = \Lambda_{i1}^{\tau} = 0$ for $i = 1, \ldots, N$ and for $\tau = 1, \ldots, 4$.[16]. The table shows that there is significant statistical evidence showing the importance of *starred* uncertainty and *starred* unemployment in each of the state-specific equations. In other words, there is evidence of substantial interaction of economic activity and behaviour across disaggregate units, and therefore supports the use of the GVAR modelling approach as an appropriate framework to capture these complex interactions.

5.2.2 Predictive Criteria

This section makes use of a statistical criterion to investigate the usefulness of disaggregate data in understanding the effect of uncertainty at the U.S state and national level. While a well-specified disaggregated model will generally outperform an aggregate model, if the

[15] The t-statistic threshold follows Clements and Hendry (2005) who find that the inclusion of variables below the threshold damages the predictive ability of the AR model. The specification search with the threshold also ensures stability for dynamic analysis when working with a large system of variables.

[16] For each bivariate system for each state, this involves for (i) $\tau = 1$, the significance *starred* uncertainty on uncertainty; (i) $\tau = 2$, the significance of *starred* unemployment on uncertainty; (iii) for $\tau = 3$, the significance of *starred* uncertainty on unemployment and (iv) $\tau = 4$, the significance of *starred* unemployment on unemployment.

disaggregate model is misspecified, this may not hold. For instance, a disaggregate model may be misspecified if macroeconomic influences are incorrectly omitted or if measurement errors found in the disaggregate model cancel out in the aggregate. The inclusion of the global variables in the GVAR model means that the first will unlikely to be a problem, however, the second might be potentially relevant due to the sample size of the Google data. The prediction criteria test proposed in PPK assesses the ability of a disaggregate model in predicting an aggregate series of interest, relative to the ability of an aggregate model, under the null that the disaggregate model is true.

In this context, assume that each state, i, is modelled according to the VARX* model specified earlier and that the aggregate model is assumed to be estimated of the following form:

$$\begin{bmatrix} US_t \\ UR_t \end{bmatrix} = \begin{bmatrix} \mathbf{c}_t^{US} \\ \mathbf{c}_t^{UR} \end{bmatrix} + \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \begin{bmatrix} US_{t-1} \\ UR_{t-1} \end{bmatrix} + \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix} \begin{bmatrix} US_{t-2} \\ UR_{t-2} \end{bmatrix} + \begin{bmatrix} e_t^{US} \\ e_t^{UR} \end{bmatrix} \tag{8}$$

where $US_t = \sum_{i=1}^{s} w_i us_{i,t}$ and $UR_t = \sum_{i=1}^{s} w_i ur_{i,t}$. In vector form, where $y_t = (US_t, UR_t)$, this can be written as:

$$y_t = c_t + \mathbf{a} y_{t-1} + \mathbf{b} y_{t-2} + \mathbf{e}_t. \tag{9}$$

In PPK, the following statistics are used to rank the disaggregate model and the aggregate model:

$$s_a^2 = \frac{\mathbf{e_t}'\mathbf{e_t}}{T - \kappa_a}, \tag{10}$$

$$s_d^2 = \sum_{i,j=1}^{N} \frac{\sum_{t=1}^{n} (w_i u_{it})' w_j u_{jt}}{T - \kappa_i - \kappa_j + tr(A_i A_j)} \tag{11}$$

where T is the number of observations;, κ_a is the number of estimated parameters in the aggregate model, κ_i is the number of unrestricted parameters in each disaggregate model; and $A_i = X_i(X_i'X_i)^{-1}X_i'$, and X_i is the matrix which consists of the explanatory variables in the i-th equation These statistics have the property that on average $s_a^2 > s_d^2$ if the disaggregate model is true. Lee and Shields (1998) transform s_a^2 and s_d^2 to statistics that are comparable to R^2 measures:

$$r_a^2 = 1 - \frac{s_a^2}{\sum_{t=1}^{T} [y_t - \overline{y_t}]^2 / [T - \kappa_a]}, \text{ and} \tag{12}$$

$$r_d^2 = 1 - \frac{s_d^2}{\sum_{t=1}^{T} [y_t - \overline{y_t}]^2 / [T - \kappa_a]} \tag{13}$$

where $\overline{y_t}$ represents the mean of the aggregate uncertainty and unemployment series and r_a and r_d denote the transformed prediction criteria for the aggregate and disaggregate model respectively. In this case, on average, r_a^2 will be smaller than r_d^2 if the disaggregate model is true. Asymptotically, if the disaggregate model is true, the criterion discussed above will rank the models correctly, on average, but this may not hold in a finite sample. To assess this, following GLP, a simulation experiment can be carried out by simulating the distribution of the test statistic, $D_{ad} = r_d^2 - r_a^2$, defining the difference between the choice criteria. Accordingly, if the disaggregate model is correctly specified, $D_{ad} > 0$. However, although this will hold asymptotically, this statistic may be smaller than zero in any particular sample simply because of sample variation.[17] Given this, GLP suggest calculating a value $d_{ad}^*(\alpha)$ obtained from the distribution of D_{ad} such that the probability of selecting the disaggregate model when it is true is $(1 - \alpha)$. We do so by first simulating the data 5000 times, assuming that the structure of the disaggregate model is the true data-generating process, and sampling with replacement. At each run, the test-statistic is recalculated to obtain the distribution for D_{ad} and the value of $d_{ad}^*(\alpha)$ is determined from the left tail of this distribution.

We find significant statistical evidence that the disaggregate model outperforms the aggregate model in terms of its ability to predict aggregate uncertainty and aggregate unemployment as shown in Figure 3. Figure 3 reports the predictive difference between the aggregate model and the disaggregate model in predicting uncertainty (top panel)

[17]In this case, it would not be strictly valid to reject the disaggregate model simply because $D_{ad} < 0$.

and unemployment (middle panel) - that is $D_{ad} = r_d^2 - r_a^2$ along $d_{ad}^*(\alpha)$ from the simulated distribution ($\alpha = 0.05$). The figure clearly shows that the disaggregate model displays a higher ability to predict uncertainty and unemployment compared to the aggregate model, as seen by positive D_{ad} values (in red). Further, once random variation is taken into account, there is still significant evidence to support the use of the disaggregate model, as illustrated by the fact the test statistics for predicting uncertainty and unemployment are both larger and on the right of the left tails of the distribution, $d_{ad}^*(\alpha)$.

In a second exercise, we specifically investigate the value-added of our state-level measures of uncertainty in the disaggregate model. The previous exercise showed that a disaggregate model with state-level uncertainty measures and the state-level unemployment outperforms the aggregate model in terms of predicting aggregate variables. We further investigate whether the same disaggregate model outperforms a partial disaggregate model, which jointly models an aggregate measure of uncertainty (replacing state-level measures of uncertainty) together with state-level measures of unemployment, in terms of predicting the aggregate unemployment variable, under the null that the fully disaggregate model is true. We find significant statistical evidence that the full disaggregate model outperforms the partial disaggregate model as shown in the bottom panel of Figure 3. This is reflected in the figure by a positive D_{dpard} value (in red). Further, once random variation is taken into account, there is still significant evidence to support the use of the disaggregate model, as illustrated by the fact the test statistics for predicting unemployment are both larger and on the right of the left tail of the distribution, $d_{dpard}^*(\alpha)$.

5.3 Dynamic Impulse Response Analysis

The dynamic analysis makes use of the generalised impulse response functions (GIRFs) framework introduced in Koop et al. (1996) and adapted to VAR/VECM models in Pesaran and Shin (1998), and GVAR in Dees et al. (2007). However, within this framework, we make use of a broad time-ordering assumption, namely that uncertainty is formed before business owners make decisions on unemployment (whilst still allowing uncertainty shocks between states to be correlated). This assumption is consistent with the short-run identification assumption typically used in the uncertainty literature – see Caggiano et al. (2014); Baker et al. (2016) or Fernández-Villaverde et al. (2015), for example. This also assigns the maximum possible effect to the uncertainty shocks and provides an upper bound on the measure of the effect of uncertainty.

Consider the GVAR(2) model in Equation 7, the moving average representation is given by:

$$x_t = \epsilon_t + A_1\epsilon_{t-1} + A_2\epsilon_{t-2} + \dots \tag{14}$$

and A_s can be defined recursively as:

$$A_s = F_1 A_{s-1} + F_2 A_{s-2}, \quad s = 1, 2, \dots \tag{15}$$

with $A_0 = I$, $A_s = 0$ for $s < 0$.

The time-ordering assumption of the shocks allows us to separate the effects of the uncertainty shocks, denoted as $\delta_t = (\epsilon_{1t}, \epsilon_{3t}, \dots, \epsilon_{101t})$, from the total shocks to the GVAR by regressing ϵ_t on δ_t and writing $\epsilon_t = \bar{D}\delta_t + \tilde{\epsilon}_t$. In this case, In this case, Equation 14 can be rewritten in terms of 'orthogonal' uncertainty shocks and other unemployment-related shocks:

$$x_t = \left[\bar{D}\delta_t + \tilde{\epsilon}_t\right] + A_1[\bar{D}\delta_{t-1} + \tilde{\epsilon}_{t-1}] + A_2[\bar{D}\delta_{t-2} + \tilde{\epsilon}_{t-2}] + \dots \tag{16}$$

The impulse response function for the GVAR model can then be written as:

$$IRF(x_t; u_{it}, h) = \frac{e_i' A_h \bar{D}\Sigma_\delta \bar{D}' e_j}{\sqrt{e_j'\Sigma_u e_j}}, h = 0, 1, 2, \dots; \; i, j = 1, 2, \dots, 2N \tag{17}$$

where Σ_δ denotes the variance-covariance matrix of the uncertainty shocks, δ_t, e_i and e_j are $2N \times 1$ selection vectors with unity in their i-th, j-th elements respectively and zeros elsewhere and h is the horizon of the impulse. For an aggregate shock, e_i is a vector of aggregate weights in the elements of $j = 2, 4, 6, ..., 2N$, in the case of a national unemployment shock for instance, where the weights sum to one and where population weights are chosen to be consistent with the construction of the *starred* variables in the

GVAR model. Given the identification assumption, the effect of a population-weighted aggregate uncertainty shock on the (weighted-aggregate) unemployment variable in the GVAR model can be seen in Figure 4. This figure also reports the traditional orthogonalised impulse response function for the bivariate VAR aggregate model showing the impact of an orthogonal uncertainty shock (defined through a Cholesky decomposition) on the unemployment variable. Although the impact effects of the respective uncertainty shocks on unemployment in both the aggregate model and the disaggregate model are similar, we find that the impact of an aggregate uncertainty shock in the disaggregate model is significantly smaller than in the aggregate model when we look at the peak response of unemployment. Since unemployment is modelled as the quarterly change in the unemployment rate, we recover the response of the actual unemployment rate to provide intuitive interpretations on the effect of uncertainty shocks.[18] We find that the peak effect of uncertainty shock in the disaggregate GVAR model causes the unemployment rate to go up by 0.1 percentage points. On the other hand, the unemployment rate goes up by 0.28 percentage points in the aggregate model at peak. These findings are within the bounds of the measures documented in the literature. For instance, Caggiano et al. (2014) find that unemployment goes up at peak by 0.4 percentage points (in bad times) or by 0.2 percentage points (in normal times) after an uncertainty shock. On the other hand, Leduc and Liu (2016) find that the peak effect of uncertainty on the unemployment rate is equal to 0.15 percentage points. Similarly, Mumtaz (2018) finds that the unemployment rate rises by 0.25%. On the whole, even after assigning uncertainty to have maximum effects due to the time-ordering assumption, our results are shown to be at the more conservative end of the impact of uncertainty on economic activity. In terms of the dynamics, both impulse responses tend to start dying away after two years although, the GVAR model, despite providing a smaller peak response effect, shows a far more prolonged process of dissemination of the initial shock taking approximately double the time of the impulse response from the VAR model which rapidly disseminates after approximately two years.

In sum, both modelling frameworks point to the significant effect of uncertainty on

[18] Calculations showing how to recover the impact on the actual unemployment rate are in the Appendix to the paper.

unemployment. The use of an aggregate model however, shown to be misspecified in the previous section, accordingly averages over relevant feedbacks and interactions through the process of aggregation in the analysis of the impact of uncertainty on the economy and therefore over-estimates the effects of uncertainty shocks relative to the disaggregate model. The aggregate model also underestimates the time it takes for the uncertainty shock to work its way through the system. As a result, the disaggregate model shows the effect of an uncertainty shock on unemployment to be statistically and significantly smaller than that according to the aggregate model and with a relatively far more prolonged dynamic response.

We close this section by observing whether there are any spatial patterns apparent in the heterogeneous responses of state-level unemployment to an aggregate uncertainty shock. In a heat map of the U.S., Figure 5 presents the median estimate of the peak response of state-level unemployment to a one standard deviation aggregate uncertainty shock. We find that unemployment increases in all states in response to an increase in U.S.-wide uncertainty. The findings are also consistent to those documented in Mumtaz et al. (2018) who find that real income declines in all states in response to an increase in U.S.-wide uncertainty (even without accommodating for state-level uncertainty). The figure also shows that the magnitude of the increase in unemployment is largest in the coastal states. On the other hand, unemployment in more central states in the U.S. seems to be less affected. These heterogeneous responses in the response of unemployment to uncertainty shocks could potentially be driven by cross-state variations in financial and fiscal conditions, the industry mix, and the labour market. We investigate this in more detail in Section 6.

5.4 Decomposition Analysis: State versus National Influences

In this section, we explore, for each state, the relative importance of national influences relative to state-specific influences of a national uncertainty shock on state-level economic activity. This section follows Garratt et al. (2018) and characterises the dynamic effects of specified shocks by using the variance-based persistence profile (PP) measure proposed by Lee and Pesaran (1993). The PP is used to measure the long-run response of the level series to shocks and trace out the accumulated response over time to characterise the system dynamics.

At time horizon h, the PP's are defined by the $2N \times 2N$ matrix $P(h)$ and as $h \to \infty$, converges to the following persistence matrix, in which the (i, j)-th element is given by:

$$p_{ij} = \frac{e_i' A(1) \Sigma_\epsilon A(1)' e_j}{\sqrt{(e_i' A(0) \Sigma_\epsilon A(0)' e_i)(e_j' A(0) \Sigma_\epsilon A(0)' e_j)}}, \qquad i, j = 1, ..., 2N, \tag{18}$$

where p_{ij} measures the infinite-horizon effect of system-wide shocks to variables in the system. For instance, the permanent impact on unemployment in each state of a system-wide shock which causes the unemployment variable in each state to rise by one standard error on impact, can be captured by considering the measures p_{ii}, where $i = 2, 4, 6, \ldots 2N$ and where $i = j$.

In this analysis, two decompositions of the persistence measure p_{ii} are of interest. The first concerns the part due to orthogonalised uncertainty shocks on state-unemployment dynamics relative to other unidentified unemployment-related shocks, and this decomposition makes use of the orthogonalisation of uncertainty shocks as described in the previous section. The second concerns the decomposition of the dynamic propagation of these uncertainty shocks into the components due to national dynamics relative to state-specific dynamics, focusing on the infinite-horizon effect on state-specific unemployment.

For the first decomposition, as described in the previous section, the assumption of a time-ordering allows us to separate uncertainty shocks, δ_t, from other unidentified unemployment-related shocks $\tilde{\epsilon}_t$, where $\epsilon_t = \bar{D}\delta_t + \tilde{\epsilon}_t$, $\Sigma_\epsilon = \bar{D}\Sigma_\delta \bar{D}' + \Sigma_{\tilde{\epsilon}}$ with Σ_δ and $\Sigma_{\tilde{\epsilon}}$ representing the variance-covariance matrices of δ_t and $\tilde{\epsilon}_t$, respectively. In this case, the numerator of the persistence measure in Equation 18 can be decomposed as:

$$\begin{aligned} e_i' A(1) \Sigma_\epsilon A(1)' e_j &= e_i' A(1) [\bar{D} \Sigma_\delta \bar{D}' + \Sigma_{\tilde{\epsilon}}] A(1)' e_j \\ &= e_i' A(1) \bar{D} \Sigma_\delta \bar{D}' A(1)' e_j + e_i' A(1) \Sigma_{\tilde{\epsilon}} A(1)' e_j \end{aligned}. \tag{19}$$

Dividing these two terms by the denominator of Equation 18 decomposes the contribution of the persistent effect on state-specific unemployment, of a system-wide shock that causes the unemployment variable in each respective state to rise by one standard error on impact, into the part due to orthogonalised uncertainty shocks, p_{ii}^U, and the remaining part due to the unidentified unemployment-related shocks, p_{ii}^V, where $p_{ii} = p_{ii}^U + p_{ii}^V$ and where p_{ii}^U can be written as:

$$p_{ii}^U = \frac{e_i' A(1) \left[\bar{D}\Sigma_\delta \bar{D}'\right] A(1)' e_i}{\sqrt{(e_i' A(0)\Sigma_\epsilon A(0)' e_i)(e_i' A(0)\Sigma_\epsilon A(0)' e_i)}}. \tag{20}$$

As mentioned earlier, this decomposition assigns the maximum possible effect to the uncertainty shocks so that the persistence measures obtained with this time-ordering assumption provide a useful upper bound on the measure of the effect of uncertainty.

For the second decomposition, noting that the influence of the national elements in the GVAR framework is captured through the *starred* parameters, the PP can be decomposed to show the dynamic effects of shocks propagating through national and state-specific elements. To see this, it is useful to separate the state-specific parameters from the national parameters by rewriting Equation 7 as:

$$x_t = (I - B_0^*)^{-1} \left(a_0 + B_1 x_{t-1} + B_1^* x_{t-1} + B_2 x_{t-2} + u_t\right) \tag{21}$$

where

$$B_l = \begin{bmatrix} \Phi_{1,l}\bar{W}_1 \\ \Phi_{2,l}\bar{W}_2 \\ \vdots \\ \Phi_{N,l}\bar{W}_N \end{bmatrix}, \; B_0^* = \begin{bmatrix} \Lambda_{1,0}W_1^* \\ \Lambda_{2,0}W_2^* \\ \vdots \\ \Lambda_{N,0}W_N^* \end{bmatrix}, \; B_1^* = \begin{bmatrix} \Lambda_{1,1}W_1^* \\ \Lambda_{2,1}W_2^* \\ \vdots \\ \Lambda_{N,1}W_N^* \end{bmatrix}, \text{ for } l = 1, 2,$$

and $\bar{W}$ contains the top k elements of the $(k_i + k_i^*) \times k$ W_i matrix while W^* denotes the bottom k^* of W_i, and the variance covariance matrix is $\Sigma_u = E(u_t u_t')$. It is straightforward to see that in the absence of national shocks, B_0^* would be zero and there would be no national dynamics if $B_1^* = 0$. To simplify the notation, and writing $(I_k - B_0^*)^{-1} = I + M^*$, the reduced-form GVAR(2) model in Equation 7 can be decomposed into state-specific components (denoted by superscript S) and the national components (denoted by superscript N) as follows:

$$x_t = (b_0^S + b_0^A) + (F_1^S + F_1^N)x_{t-1} + (F_2^S + F_2^N)x_{t-2} + \epsilon_t \quad \text{for } t = 1, \ldots T \tag{22}$$

where $b_0^S = a_0$; $F_l^S = B_l$ for $l = 1, 2$ and $b_0^N = M^* a_0$; $F_l^N = M^* B_l + (I + M^*)B_l^*$ for $l = 1, 2$. Then, following Garratt et al. (2018), elements of $A(L)$ in the moving average representation in Equation 14 can be decomposed into a national element $A^N(L)$, capturing the influence of national dynamics, and into a state-specific element $A^S(L)$, which abstracts completely from the influence of the *starred* parameters:

$$x_t = \epsilon_t + (A_1^S + A_1^N)\epsilon_{t-1} + \ldots \quad \text{for } t = 1, \ldots, T, \tag{23}$$

where $A_1^S = F_1^S A_{s-1}^S + F_2^S A_{s-2}^S$ for $s = 1, 2$ with $A_0^S = I$ and $A_k^S = 0$ for $k < 0$. Note that the national effects are captured by the difference between total effects and state-specific effects, $A_k^N = A_k - A_k^S$ for $k < 0$.

The elements of the infinite-horizon persistence measure matrix in Equation 18 can then be rewritten as:

$$p_{ij} = \frac{e_i' \left[A^S(1) + A^N(1)\right] \Sigma_\epsilon \left[A^S(1) + A^N(1)\right]' e_j}{\sqrt{(e_i' A(0) \Sigma_\epsilon A(0)' e_i)(e_j' A(0) \Sigma_\epsilon A(0)' e_j)}}. \tag{24}$$

It follows from Equation 20 that the influence of orthogonalised uncertainty shocks propagating purely through state-specific dynamics, as defined by the elements of $A^S(1)$, on the unemployment variable in each state, can be isolated through the following expression:

$$p_{ii}^{SU} = \frac{e_i' A^S(1) \left[\bar{D} \Sigma_\delta \bar{D}'\right] A^S(1)' e_i}{\sqrt{(e_i' A(0) \Sigma_\epsilon A(0)' e_i)(e_i' A(0) \Sigma_\epsilon A(0)' e_i)}},$$

and where $p_{ii}^{NU} = p_{ii}^U - p_{ii}^{SU}$, defined as the remainder term, gives the importance of national dynamics in propagating uncertainty shocks in the impact on state-specific unemployment.

Figure 6 reports the ratio: $\frac{p_{ii}^{NU}}{p_{ii}^U}$ in a heat map of the U.S., which gives the relative importance of national dynamics versus both state and national dynamics, as modelled in Equation 7, in terms of respective measures of the persistent effect of the above-defined uncertainty shocks on state-specific unemployment. The persistence measures are scaled by a standard error state-specific unemployment shock on impact. A simple correlation between the statistics for each state represented in Figure 5 and those in Figure 6 gives a value of 0.4 showing that national influences play a clear role in propagating the effects of uncertainty. On average, the uncertainty shocks transmitted through the channels associated with national elements account for 37% of the total variation in state-specific unemployment and the remaining 53% is attributed to the purely state-specific channel abstracting entirely from national dynamics. However, the figure also shows that there is a significant degree of heterogeneity across the U.S. states. The figure also shows that the more central states in the U.S. seem to be less affected by national dynamics in the influence of uncertainty on economic activity - such as in the states of Texas, New Mexico or Colorado, for instance. On the other hand, unemployment in the eastern states, tend to be far more influenced by national elements in the propagation of uncertainty shocks. Factors that might explain such heterogeneity could include state-specific characteristics such as the state-level industry composition, fiscal constraints, labour market constraints and financial frictions. We explore these patterns in terms of potential pointers towards an economic narrative in more detail in the next section.

6 An Exploration into the Heterogeneity of State-Level Responses and National Influences

This section uses state-specific characteristics for an investigation into factors potentially important when considering (i) the heterogeneous responses of state-level unemployment to a U.S. wide uncertainty shock as detailed in Section 5.3; and (ii) the heterogeneous importance of national influences in the role of the propagation of uncertainty shocks on state-level unemployment as described in Section 5.4. The analysis follows the regression specifications employed in Mumtaz et al. (2018):

$$\rho_i = c + \beta X_i + R_i + \eta_i, \qquad\qquad i = 1, ..., N \tag{25}$$

where the dependent variable is given respectively by (i) $\rho_i = response_i$ reflecting the peak response of state i's unemployment to a U.S-wide uncertainty shock when considering the heterogeneous responses of unemployment as plotted in Figure 5; and (ii) $\rho_i = \frac{p_{ii}^{NU}}{p_{ii}^{U}}$ when considering the relative importance of national influences in propagating uncertainty shocks on state i's unemployment; c denotes the intercept; X_i include the explanatory variables depicting state-specific characteristics; and R_i denote regional dummies according to the U.S Bureau of Economic Analysis. A description of the data sources can be found in the Appendix.

The testing strategy follows that employed in Mumtaz et al. (2018) and involves specifying explanatory variables capturing various sets of influences and investigating their respective explanatory power through sequential regressions. The regression specifications and testing strategy are described below. In terms of the explanatory variables, we specify five sets of regressors. The first set represents the industry structure for each state and contain information on the state-level GDP share of the Agriculture, Construction, Finance, Government, Manufacturing, Mining, Real Estate. The second set accounts for the degree of financial frictions. Following Carlino and DeFina (1998), financial frictions are defined as the percentage of each state's loans made by small banks.[19] In addition to small banks' loans, the home ownership rate is also included to account for cross-state differences in the housing market. The third set measures the labour market frictions by using data on the percentage of workers represented by unions for each state and the right-to-work (RTW), which is a dummy variable for whether a state has 'right-to-work' legislation as of 2016. The fourth set captures the level of the fiscal condition by using data on the state-level debt-GDP ratio, intergovernmental transfers and the state-level spending-GDP ratio. Finally, the fifth set considers the share of internet users across different U.S states and serves as a robustness check given our internet-data-based measure

[19] Smalls banks are banks whose assets are less than the 90 % of the national average.

of uncertainty.

6.1 Heterogeneity of State-Level Responses of Unemployment

The regression strategy involves starting with an initial regression of $\rho_i = response_i$ and solely considering the first set of explanatory variables which includes variables representing state industry structures and relevant variables. Starting with just one industry variable, and retaining only if statistically significant (at the 10% level of significance), other industry variables are introduced sequentially in subsequent regressions, once again retaining significant variables in the regression. In the following regressions, additional control variables from the other sets of regressors are introduced one by one in sequential regressions and only retained if statistically significant. Table 4 reports the coefficients from these sequential regressions.

Table 4 Column 1 presents our benchmark regression result from the first sequence of regressions relating the estimated responses ρ_i to variables accounting for the structure of industry in each state where only the remaining significant industry variables are reported.[20] The column shows that only the share of the manufacturing and the mining industry have a robust and significant relationship with the responses of unemployment rate to an increase in U.S.-wide uncertainty. These findings on the industry mix are completely consistent with those reported in Mumtaz et al. (2018). We find that states with a higher concentration of the manufacturing industry experience a larger increase in unemployment after an increase in U.S.-wide uncertainty. On the other hand, we find that states with a larger mining industry experience a relatively smaller increase in unemployment. As documented in many studies such as Jo (2014), Mumtaz et al. (2018) or Tran (2019), for instance, there is evidence that an uncertainty shock results in an increase in the price of some commodities. If this is the case, then this could be a contributing factor in negating the negative effects of uncertainty in states with a higher mining share of industry.

Column 2 and Column 3 of Table 4 show that the variables reflecting the housing

[20]Details of the regressions underlying this benchmark result are reported in the Appendix.

market (Home Vacancy or Home Ownership) are not significant and are thus excluded from the subsequent regression specifications. Columns 4 – 10 present the results of sequentially including variables reflecting: labour market conditions (i.e. Right-to-work or Union Membership); financial conditions; and the fiscal condition (state-level debt, state-level spending and intergovernmental transfers). We find that none of these variables have a statistically significant relationship with the responses of unemployment to state-level uncertainty – with Manufacturing and Mining still remaining significant in most of these regressions. There is good evidence therefore to support the results reported in Column 1. Findings from these subsequent regressions (Column 2 onwards) however are in contrast to the results in Mumtaz et al. (2018) who report a significant association between the effect of uncertainty and the state-level housing market, the fiscal condition and the financial constraint - although with using real income as a measure of real economic activity. When we use unemployment as a measure of real economic activity and in a framework accommodating state interactions and feedbacks, the relationship of these variables with uncertainty shocks is weaker.

In summary, the benchmark estimates and the detailed robustness checks suggest the following results. First, states with a higher concentration of the manufacturing industry experience a larger increase in unemployment when impacted by an uncertainty shock. Second, states with a higher concentration of the mining industry appear to be affected less by this shock. Last, while the first two results are similar to the findings documented by Mumtaz et al. (2018), we do not find any significant association between the share of the construction industry, the state-level fiscal position, the labour market condition and the financial constraint in terms of the response of state-level unemployment.

6.2 Heterogeneity of National Influences in the Propagation of Uncertainty Shocks

In this section, we employ the same regression and testing strategy as described in the previous section, with the dependent variable now $\rho_i = \frac{p_{ii}^{NU}}{p_{ii}^{U}}$. Again, starting with a set of sequential regressions involving the first set of variables reflecting the industry structure (and as detailed in the Appendix), we end up with an industry mix involving Government and Real Estate remaining as the significant variables as the initial baseline regression and as reported in Column 1 of Table 5. To this baseline regression, we add Home Vacancy in Column 2 and Home Ownership in Column 3. Given the statistical significance of Home Vacancy, we keep it in the regression specification and add in sequentially Right to Work in Column 4, Union membership in Column 5 and the degree of financial frictions in Column 6. None of these latter three variables prove to be statistically significant. Home Vacancy also drops out with the inclusion of the measure of state-level debt.[21] We subsequently add in a measure of intergovernmental transfers as reported in Column 9. Although we find that the measure of intergovernmental transfers is statistically significant, we exclude this variable from the regression on the grounds that it takes on a value of almost zero, reduces the overall goodness-of-fit of the model as seen by the adjusted R^2 measure in Columns 8 – 10, and its inclusion takes away from the significance of the share of the government sector which has been a robust explanatory variable in all the other regressions.

Our preferred specification is as in Column 8. The results in this column shows that only the government sector, the real estate industry and the state-level of fiscal conditions are robust in terms of statistical significance in terms of the association the relative importance of the national influences in propagating uncertainty shocks. We find that states with a higher concentration of the real estate industry reflect a tendency for a state to experience a larger impact through national influences in propagating uncertainty shocks. This could quite possibly relate to the impact of the recent Great Financial Crisis through the origins of the housing bubble and the subprime mortgage lending crisis - a national phenomena.

On the other hand, states with a more active fiscal policy experience less of an influence from national factors in the propagation of uncertainty shocks. We find a negative and statistically significant relationship between the importance of national influences in propagating uncertainty shocks and the share of the government sector and the size of

[21] The increase in the goodness-of-fit of the model is also another reason for keeping state-level debt and removing the Home Vacancy variable. We can see that the adjusted R^2 measure increases when we replace Home Vacancy by state-level debt in 5 Column 3 and Column 8.

the state-level debt. Our conjecture on the negative association between an active fiscal policy and the importance of national influences in propagating uncertainty shocks is further solidified when we replace state-level debt by state-level spending in Table 5 Column 11. We find that there is a significant association between state-level spending and the importance of national factors in the propagation of uncertainty shocks. Finally, the inclusion of the share of internet users in Column 12 does not affect the results of our preferred specification.

In summary, the results from our preferred regression specification and detailed robustness checks suggest the following: the relative importance of national factors in the propagation of uncertainty shocks tends to be (i) smaller in states with a more active fiscal policy; and (ii) greater in states with a larger concentration of the real estate industry.

7 Conclusion

Uncertainty is commonly identified as being a major driver in causing the Great Recession and the subsequent slow recovery. Whilst most studies focus on studying the impact of uncertainty at the aggregate level, this paper estimates the impact of uncertainty shocks in a disaggregate model featuring state-level economic activity and uncertainty proxies, where Google search data is used as proxies for uncertainty. The disaggregate framework of analysis is in the context of the GVAR model which is able to characterise the evolution of state-specific variables whilst incorporating interactions and spillover effects across states.

We find there is strong evidence to support the usefulness of disaggregated data in models of uncertainty and economic activity for the U.S., based on the ability of the disaggregate model to predict aggregate uncertainty and unemployment. The use of a disaggregate model accommodates and captures relevant feedbacks and interactions in the analysis on the impact of uncertainty on the economy - which the misspecified aggregate model omits. As a result, the disaggregate model exhibits significantly smaller effects of uncertainty shocks on the economy with a far more prolonged dissemination of these shocks compared to the aggregate model.

There is a large degree of heterogeneity across states in the respective responses of unemployment to a national uncertainty shock and also in the respective importance of national influences in propagating uncertainty shocks. Cross-sectional regression results suggest that the responses of unemployment are positively associated with the share of the manufacturing industry and negatively associated with the concentration of the mining industry. We also find that national influences are relatively less important than the state-specific influences in propagating the effect of uncertainty shocks and found to be positively associated with the size of the real estate industry and negatively associated with states exhibiting a more active fiscal policy, a greater share of the government sector, and higher levels of debt and spending. Further analysis on providing an economic commentary on the mechanisms underlying these results is certainly warranted and could provide important insights in informing state and national policies to mitigate the negative effects of uncertainty on the economy.

References

Angelini, G., Bacchiocchi, E., Caggiano, G., and Fanelli, L. (2018). Uncertainty across volatility regimes. *Journal of Applied Econometrics*, forthcoming.

Bachmann, R., Elstner, S., and Sims, E. R. (2013). Uncertainty and economic activity: Evidence from business survey data. *American Economic Journal: Macroeconomics*, 5(2):217–249.

Baker, S. R., Bloom, N., and Davis, S. J. (2016). Measuring economic policy uncertainty. *The Quarterly Journal of Economics*, 131(4):1593–1636.

Bernanke, B. S. (1983). Irreversibility, uncertainty, and cyclical investment. *The Quarterly Journal of Economics*, 98(1):85–106.

Bloom, N. (2009). The impact of uncertainty shocks. *Econometrica*, 77(3):623–685.

Bloom, N. (2014). Fluctuations in uncertainty. *Journal of Economic Perspectives*, 28(2):153–176.

Bontempi, M. E., Golinelli, R., and Squadrani, M. (2016). A new index of uncertainty based on internet searches: A friend or foe of other indicators?

Caggiano, G., Castelnuovo, E., and Groshenny, N. (2014). Uncertainty shocks and unemployment dynamics in US recessions. *Journal of Monetary Economics*, 67:78–92.

Carlino, G. and DeFina, R. (1998). The differential regional effects of monetary policy. *Review of Economics and Statistics*, 80(4):572–587.

Carriero, A., Clark, T. E., and Marcellino, M. (2018). Measuring uncertainty and its impact on the economy. *The Review of Economics and Statistics*, 100(5):799–815.

Castelnuovo, E. and Tran, T. D. (2017). Google it up! a Google Trends-based uncertainty index for the United States and Australia. *Economics Letters*.

Cesa-Bianchi, A., Pesaran, M., and Rebucci, A. (2014). Uncertainty and economic activity: A global perspective. *CAFE Research Paper*, (14.03).

Choi, H. and Varian, H. (2009). Predicting initial claims for unemployment benefits. *Google Inc*, pages 1–5.

Chudik, A. and Pesaran, M. H. (2016). Theory and practice of GVAR modelling. *Journal of Economic Surveys*, 30(1):165–197.

Clements, M. P. and Hendry, D. F. (2005). Guest editors' introduction: Information in economic forecasting. *Oxford Bulletin of Economics and Statistics*, 67(s1):713–753.

D'Amuri, F. and Marcucci, J. (2017). The power of Google searches in forecasting unemployment. *International Journal of Forecasting*, 33(4):801–816.

Dees, S., Mauro, F. d., Pesaran, M. H., and Smith, L. V. (2007). Exploring the international linkages of the euro area: a global VAR analysis. *Journal of Applied Econometrics*, 22(1):1–38.

Fernández-Villaverde, J., Guerrón-Quintana, P., Kuester, K., and Rubio-Ramírez, J. (2015). Fiscal volatility shocks and economic activity. *American Economic Review*, 105(11):3352–84.

Fernández-Villaverde, J., Guerrón-Quintana, P., Rubio-Ramírez, J. F., and Uribe, M. (2011). Risk matters: The real effects of volatility shocks. *American Economic Review*, 101(6):2530–61.

Garratt, A., Lee, K., and Shields, K. (2018). The role of uncertainty, sentiment and cross-country interactions in G7 output dynamics. *Canadian Journal of Economics*, 51(2):391–418.

Garratt, A., Robertson, D., and Wright, S. (2006). Permanent vs transitory components and economic fundamentals. *Journal of Applied Econometrics*, 21(4):521–542.

Ginsberg, J., Mohebbi, M. H., Patel, R. S., Brammer, L., Smolinski, M. S., and Brilliant, L. (2009). Detecting influenza epidemics using search engine query data. *Nature*, 457(7232):1012–1014.

Ilut, C. L. and Schneider, M. (2014). Ambiguous business cycles. *American Economic Review*, 104(8):2368–99.

Istrefi, K. and Mouabbi, S. (2017). Subjective interest rate uncertainty and the macroeconomy: A cross-country analysis. *Journal of International Money and Finance*.

Jo, S. (2014). The effects of oil price uncertainty on global real economic activity. *Journal of Money, Credit and Banking*, 46(6):1113–1135.

Jurado, K., Ludvigson, S. C., and Ng, S. (2015). Measuring uncertainty. *The American Economic Review*, 105(3):1177–1216.

Koop, G., Pesaran, M. H., and Potter, S. M. (1996). Impulse response analysis in nonlinear multivariate models. *Journal of econometrics*, 74(1):119–147.

Leduc, S. and Liu, Z. (2016). Uncertainty shocks are aggregate demand shocks. *Journal of Monetary Economics*, 82(Supplement C):20 – 35.

Lee, K. and Shields, K. (1998). Modelling sectoral output growth in the EU economies. Report for the European Commission.

Lee, K. C. and Pesaran, M. H. (1993). Persistence profiles and business cycle fluctuations in a disaggregated model of uk output growth. *Ricerche Economiche*, 47(3):293–322.

Ludvigson, S. C., Ma, S., and Ng, S. (2018). Uncertainty and business cycles: Exogenous impulse or endogenous response? New York University and Columbia University, mimeo.

Mumtaz, H. (2018). Does uncertainty affect real activity? evidence from state-level data. *Economics Letters*, 167:127–130.

Mumtaz, H., Sunder-Plassmann, L., and Theophilopoulou, A. (2018). The state-level impact of uncertainty shocks. *Journal of Money, Credit and Banking*.

Pesaran, H. H. and Shin, Y. (1998). Generalized impulse response analysis in linear multivariate models. *Economics letters*, 58(1):17–29.

Pesaran, M. H., Pierse, R. G., Kumar, M. S., et al. (1989). Econometric analysis of aggregation in the context of linear prediction models. *Econometrica*, 57(4):861–888.

Pesaran, M. H., Pierse, R. G., and Lee, K. C. (1994). Choice between disaggregate and aggregate specifications estimated by instrumental variables methods. *Journal of Business & Economic Statistics*, 12(1):11–21.

Rossi, B. and Sekhposyan, T. (2015). Macroeconomic uncertainty indices based on nowcast and forecast error distributions. *The American Economic Review*, 105(5):650–655.

Scotti, C. (2016). Surprise and uncertainty indexes: Real-time aggregation of real-activity macro-surprises. *Journal of Monetary Economics*, 82:1–19.

Shields, K., Olekalns, N., Henry, Ó. T., and Brooks, C. (2005). Measuring the response of macroeconomic uncertainty to shocks. *Review of Economics and Statistics*, 87(2):362–370.

Shoag, D. and Veuger, S. (2016). Uncertainty and the geography of the Great Recession. *Journal of Monetary Economics*, 84:84–93.

Tran, T. D. (2019). Commodity uncertainty in a small open economy: Theory and empirics'. University of Melbourne, mimeo.

Van Garderen, K. J., Lee, K., and Pesaran, M. H. (2000). Cross-sectional aggregation of non-linear models. *Journal of Econometrics*, 95(2):285–331.

Vansteenkiste, I. (2007). Regional housing market spillovers in the US: lessons from regional divergences in a common monetary policy setting.

Table 1: List of Search Terms

Affirmative action	health care act
American Rec and Reinvestment Act of 2009	health care reform
At-will employment	home price
austerity	home sales
bank loan	inflation
Bank of England	job security
Bank rate	Military budget
Bank regulation	minimum wage
bankruptcy	Monetary policy
budget cut	Money supply
business outlook	National debt of the United States
Carbon tax	National Labor Relations Act
Clean Water Act	National Labor Relations Board
collective agreement	natural reserve
Commodity Futures Trading Commission	Open market operation
construction permit	Patient Protection and Affordable Care Act
consumer confidence	pollution control
Consumer price index	price level
debt ceiling	Quantitative easing
default	real estate bubble
Discount window	recession
Dodd-Frank	reform
economic outlook	regulation
emission trading clean air act	Right-to-work law
energy policy	Securities and Exchange Commission
environment protection	Share price
Equal Employment Opportunity Commission	slow economic recovery
Equal opportunity employment	stock exchange
European debt crisis	stock market
Federal Deposit Insurance Corporation	tax cut
Federal funds rate	Tort reform
Federal Reserve System	unemployment benefit
financial crisis	unemployment extension
financial reform	United States Congress
fiscal cliff	United States Environmental Protection Agency
Fiscal policy	United States federal budget
Food and Drug Administration	United States housing bubble
food price	White House
fuel price	Workers' compensation
gas price	

Table 2: List of U.S. States

Abbreviation	State Name	Abbreviation	State Name
al	Alabama	mt	Montana
ak	Alaska	ne	Nebraska
az	Arizona	nv	Nevada
ar	Arkansas	nh	New Hampshire
ca	California	nj	New Jersey
co	Colorado	nm	New Mexico
ct	Connecticut	ny	New York
de	Delaware	nc	North Carolina
dc	District Of Columbia	nd	North Dakota
fl	Florida	oh	Ohio
ga	Georgia	ok	Oklahoma
hi	Hawaii	or	Oregon
id	Idaho	pa	Pennsylvania
il	Illinois	ri	Rhode Island
in	Indiana	sc	South Carolina
ia	Iowa	sd	South Dakota
ks	Kansas	tn	Tennessee
ky	Kentucky	tx	Texas
la	Louisiana	ut	Utah
me	Maine	vt	Vermont
md	Maryland	va	Virginia
ma	Massachusetts	wa	Washington
mi	Michigan	wv	West Virginia
mn	Minnesota	wi	Wisconsin
ms	Mississippi	wy	Wyoming
mo	Missouri		

Table 3: F-test statistics on the Joint Significance of National Uncertainty (us^*) and National Unemployment (ur^*)

	al	**ak**	**az**	**ar**	**ca**	**co**	**ct**	**de**	**dc**	**fl**	**ga**
us_{it}: sig of $us^*_{it}, us^*_{i,t-1}$	125.12***	31.6***	209.82***	27.35***	71.54***	219.42***	69.19***	4.41**	15.84***	147.04***	174.88***
us_{it}: sig of $ur^*_{it}, ur^*_{i,t-1}$	2.88*	3.2**	0.03	0.5	2.19	1.9	0.18	1.03	0.05	0.81	0.9
ur_{it}: sig of $us^*_{it}, us^*_{i,t-1}$	0.6	3.41**	0.93	2.29	1.63	1.05	1.09	1.81	0.67	0.12	5.98***
ur_{it}: sig of $ur^*_{it}, ur^*_{i,t-1}$	27.77***	12.32***	7.25***	26.9***	55.71***	11.18***	11.67***	24.48***	1.55	52.32***	37.48***
	hi	**id**	**il**	**in**	**ia**	**ks**	**ky**	**la**	**me**	**md**	**ma**
us_{it}: sig of $us^*_{it}, us^*_{i,t-1}$	55.28***	31.2***	84.17***	70.34***	21.55***	23.5***	29.67***	46.13***	19.03***	51.54***	52.74***
us_{it}: sig of $ur^*_{it}, ur^*_{i,t-1}$	1.24	0.57	0.16	0.08	0.28	0.19	5.23***	3.55**	1.31	0.58	0.21
ur_{it}: sig of $us^*_{it}, us^*_{i,t-1}$	0.15	0.1	2.45*	2.06	2.01	2.78*	0.42	2.46*	2.15	2.02	0.74
ur_{it}: sig of $ur^*_{it}, ur^*_{i,t-1}$	11.56***	11.37***	27.24***	43.16***	59.77***	45.78***	51.11***	1.68	22.45***	53.18***	11.92***
	mi	**mn**	**ms**	**mo**	**mt**	**ne**	**nv**	**nh**	**nj**	**nm**	**ny**
us_{it}: sig of $us^*_{it}, us^*_{i,t-1}$	115.31***	97.51***	47.6***	164.52***	46.46***	69.16***	10.58***	33.56***	127.88***	56.83***	184.34***
us_{it}: sig of $ur^*_{it}, ur^*_{i,t-1}$	0.17	2.61*	0.89	0.17	4.36**	0.85	0.67	0.26	0.97	2.52*	0.38
ur_{it}: sig of $us^*_{it}, us^*_{i,t-1}$	0.13	4.1**	0	1.05	1.37	0.1	0.94	0.46	0.7	1.52	0.72
ur_{it}: sig of $ur^*_{it}, ur^*_{i,t-1}$	46.79***	31.43***	17.16***	23.59***	14.02***	25.99***	28.62***	51.45***	28.56***	12.67***	40.65***
	nc	**nd**	**oh**	**ok**	**or**	**pa**	**ri**	**sc**	**sd**	**tn**	**tx**
us_{it}: sig of $us^*_{it}, us^*_{i,t-1}$	176.92***	20.44***	247.25***	31.4***	41.89***	263.65***	29.36***	103.18***	31.31***	164.03***	151.01***
us_{it}: sig of $ur^*_{it}, ur^*_{i,t-1}$	0.57	1.12	0.47	0.08	0.42	0.73	0.68	2.6*	0.2	1.44	1.12
ur_{it}: sig of $us^*_{it}, us^*_{i,t-1}$	3.32**	0.07	0.23	1.41	1.24	0.82	1.91	1.25	1.51	0	1.46
ur_{it}: sig of $ur^*_{it}, ur^*_{i,t-1}$	29.3***	24.64***	45.72***	31.71***	33.58***	48.28***	9.42***	44.69***	27.41***	42.36***	18***
	ut	**vt**	**va**	**wa**	**wv**	**wi**	**wy**				
us_{it}: sig of $us^*_{it}, us^*_{i,t-1}$	126.2***	1.35	19.12***	63.51***	17.54***	140.32***	13.31***				
us_{it}: sig of $ur^*_{it}, ur^*_{i,t-1}$	1.42	3.76**	3.15**	0.3	2.12	3.11**	2.23				
ur_{it}: sig of $us^*_{it}, us^*_{i,t-1}$	0.04	0.29	1.47	0.63	1.29	0.01	2.88*				
ur_{it}: sig of $ur^*_{it}, ur^*_{i,t-1}$	19.01***	48.44***	72.9***	21.14***	11.38***	48.19***	20.01***				

Notes: This shows the F-test statistics of the joint significance of national uncertainty ($us*$) and national unemployment ($ur*$) of the state-level uncertainty and unemployment equations $^{***}p < 0.01, ^{**}p < 0.05, ^{*}p < 0.1$.

Table 4: Heterogeneity Analysis on the Response of State-Level Unemployment to an Aggregate Uncertainty Shock

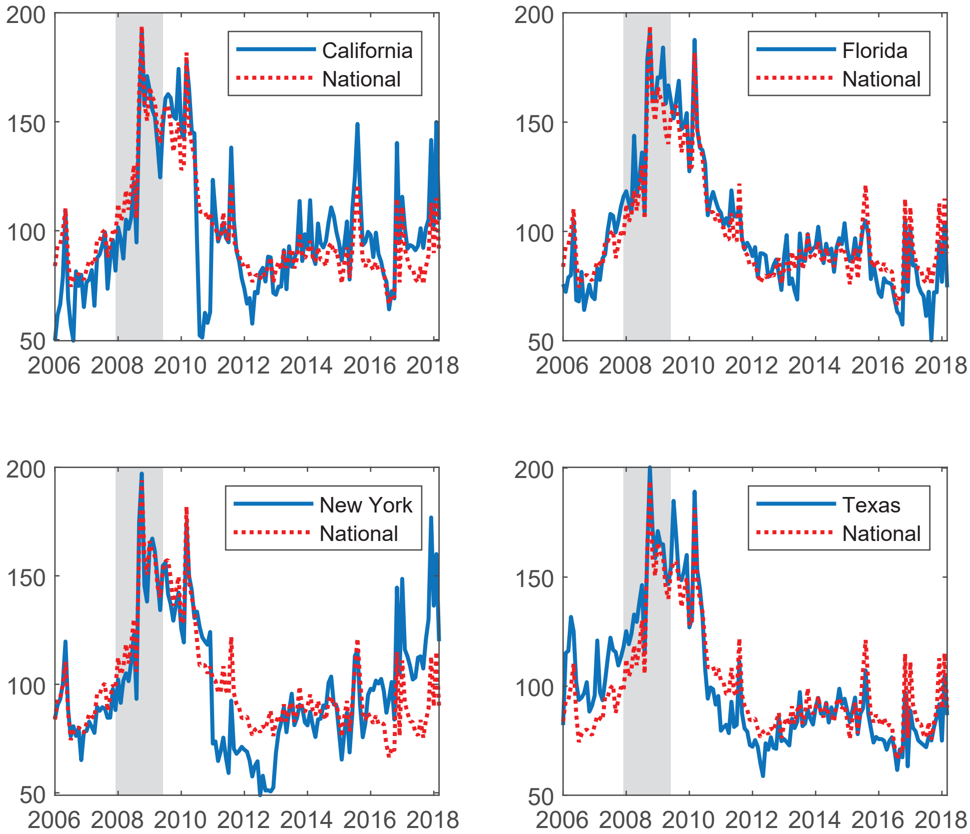
VARIABLES	(1) $response_i$	(2) $response_i$	(3) $response_i$	(4) $response_i$	(5) $response_i$	(6) $response_i$	(7) $response_i$	(8) $response_i$	(9) $response_i$	(10) $response_i$
Manufacturing	0.00152**	0.00175***	0.00135	0.00143*	0.00152*	0.00154**	0.00149*	0.00150**	0.00157**	0.00153*
	(0.000749)	(0.000611)	(0.000862)	(0.000802)	(0.000778)	(0.000753)	(0.000770)	(0.000735)	(0.000721)	(0.000763)
Mining	-0.00164*	-0.00123	-0.00173*	-0.00171*	-0.00164*	-0.00151	-0.00160*	-0.00160*	-0.00136*	-0.00165*
	(0.000905)	(0.000887)	(0.000901)	(0.000926)	(0.000894)	(0.000991)	(0.000849)	(0.000904)	(0.000769)	(0.000913)
HomeVacancy		0.0213								
		(0.0129)								
HomeOwnership			0.000417							
			(0.000763)							
RTW				0.00746						
				(0.0153)						
Union					4.28e-05					
					(0.00136)					
FinancialFriction						-0.000150				
						(0.000127)				
StateDebt							-0.000444			
							(0.00199)			
IntGovt								1.59e-10		
								(2.71e-10)		
StateSpend									-0.00149	
									(0.00165)	
InternetUsers										-1.29e-05
										(1.22e-05)
Observations	51	51	51	51	51	51	51	51	51	51
Adjusted R-squared	0.469	0.504	0.459	0.461	0.455	0.464	0.457	0.458	0.470	0.457

Notes: $response_i$ is defined in Section 5.3 and represents the peak state-level response of unemployment to an aggregate uncertainty shock. Robust standard errors in parentheses $^{***}p < 0.01, ^{**}p < 0.05, ^{*}p < 0.1$.

Table 5: Heterogeneity Analysis on the Importance of National Influences in Propagating Uncertainty Shocks

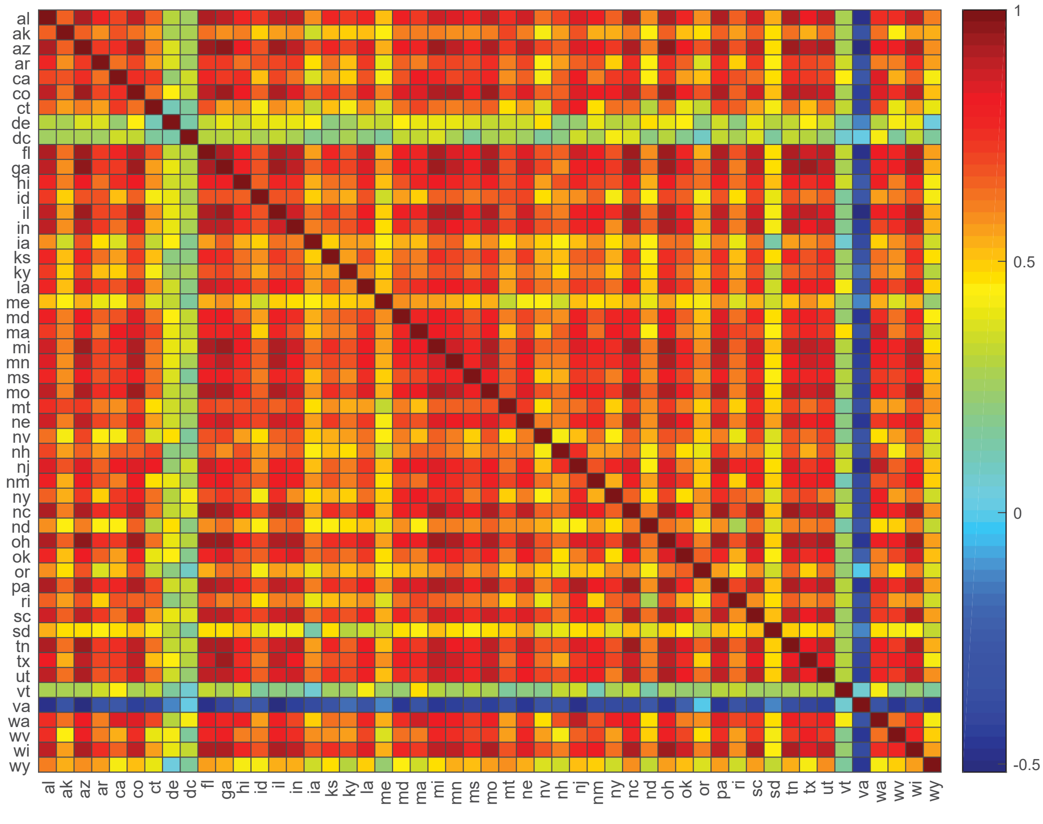
VARIABLES	(1) $\frac{p_{ii}^{NU}}{p_{ii}^{U}}$	(2) $\frac{p_{ii}^{NU}}{p_{ii}^{U}}$	(3) $\frac{p_{ii}^{NU}}{p_{ii}^{U}}$	(4) $\frac{p_{ii}^{NU}}{p_{ii}^{U}}$	(5) $\frac{p_{ii}^{NU}}{p_{ii}^{U}}$	(6) $\frac{p_{ii}^{NU}}{p_{ii}^{U}}$	(7) $\frac{p_{ii}^{NU}}{p_{ii}^{U}}$	(8) $\frac{p_{ii}^{NU}}{p_{ii}^{U}}$	(9) $\frac{p_{ii}^{NU}}{p_{ii}^{U}}$	(10) $\frac{p_{ii}^{NU}}{p_{ii}^{U}}$	(11) $\frac{p_{ii}^{NU}}{p_{ii}^{U}}$	(12) $\frac{p_{ii}^{NU}}{p_{ii}^{U}}$
Government	-1.016**	-1.036**	-0.942*	-0.965*	-1.042**	-1.146**	-1.063*	-1.062*	-0.878		-1.003*	-1.035*
	(0.420)	(0.446)	(0.552)	(0.477)	(0.460)	(0.500)	(0.564)	(0.560)	(0.591)		(0.541)	(0.569)
RealEstate	3.655***	3.626***	3.668***	3.762***	3.678***	3.706***	3.893***	3.975***	3.851***	4.045***	3.571***	3.983***
	(0.947)	(0.883)	(0.953)	(0.905)	(0.898)	(0.875)	(0.865)	(0.875)	(0.907)	(1.036)	(0.910)	(0.886)
HomeVacancy		11.77**		11.23*	10.96*	11.51**	6.100					
		(5.339)		(5.609)	(6.045)	(5.562)	(5.793)					
HomeOwnership			0.158									
			(0.451)									
RTW				4.250								
				(5.600)								
Union					-0.323							
					(0.634)							
FinancialFriction						0.0527						
						(0.0897)						
StateDebt							-1.592**	-2.018***	-1.954***	-1.868**		-1.934***
							(0.719)	(0.665)	(0.667)	(0.924)		(0.682)
IntGovt									1.94e-07*	3.27e-07**		
									(1.14e-07)	(1.45e-07)		
StateSpend											-1.136*	
											(0.644)	
InternetUsers												0.00962
												(0.00835)
Observations	51	51	51	51	51	51	51	51	51	51	51	51
Adjusted R-squared	0.508	0.548	0.497	0.542	0.539	0.539	0.574	0.574	0.572	0.552	0.525	0.565

Notes: $\frac{p_{ii}^{NU}}{p_{ii}^{U}}$ is defined in Section 5.4 and represents the relative importance of national uncertainty shocks in propagating uncertainty shocks. Robust standard errors in parentheses $^{***}p < 0.01, ^{**}p < 0.05, ^{*}p < 0.1$.

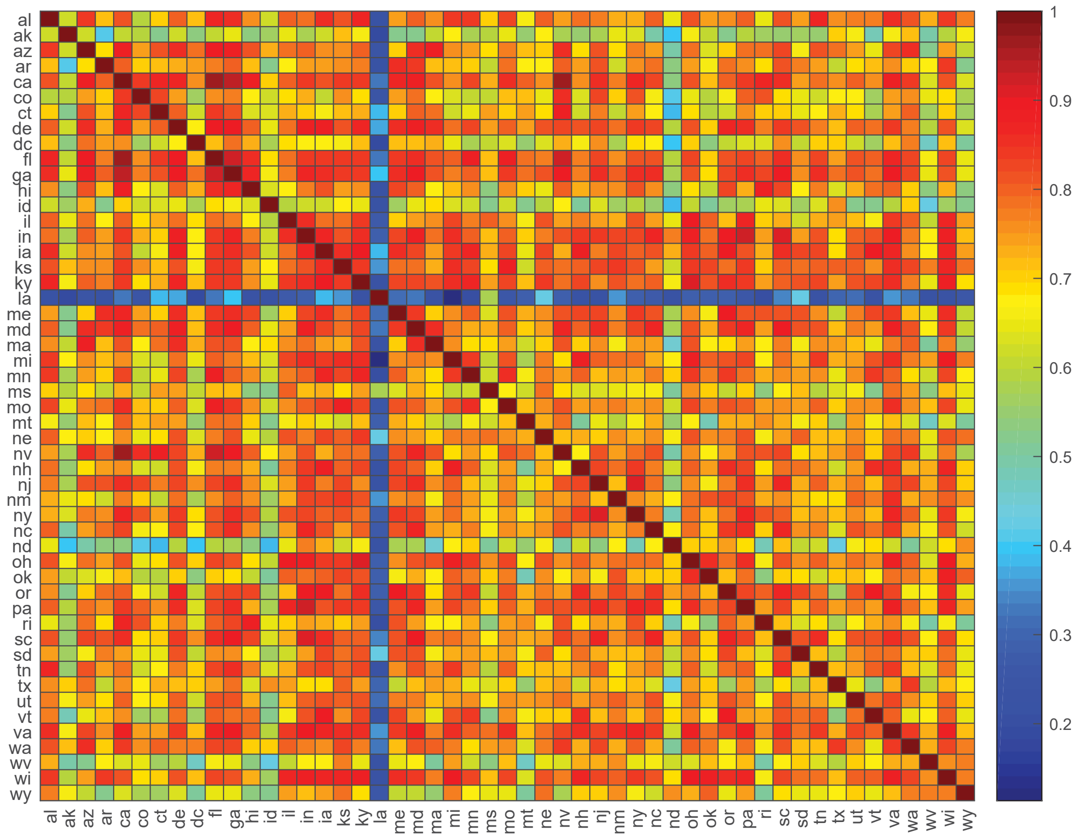


Notes: The index is transformed to have a mean of 100 and standard deviation of 30. The shaded area represents the Great Recession.

Figure 1: GTU Index for Selected States in the U.S.

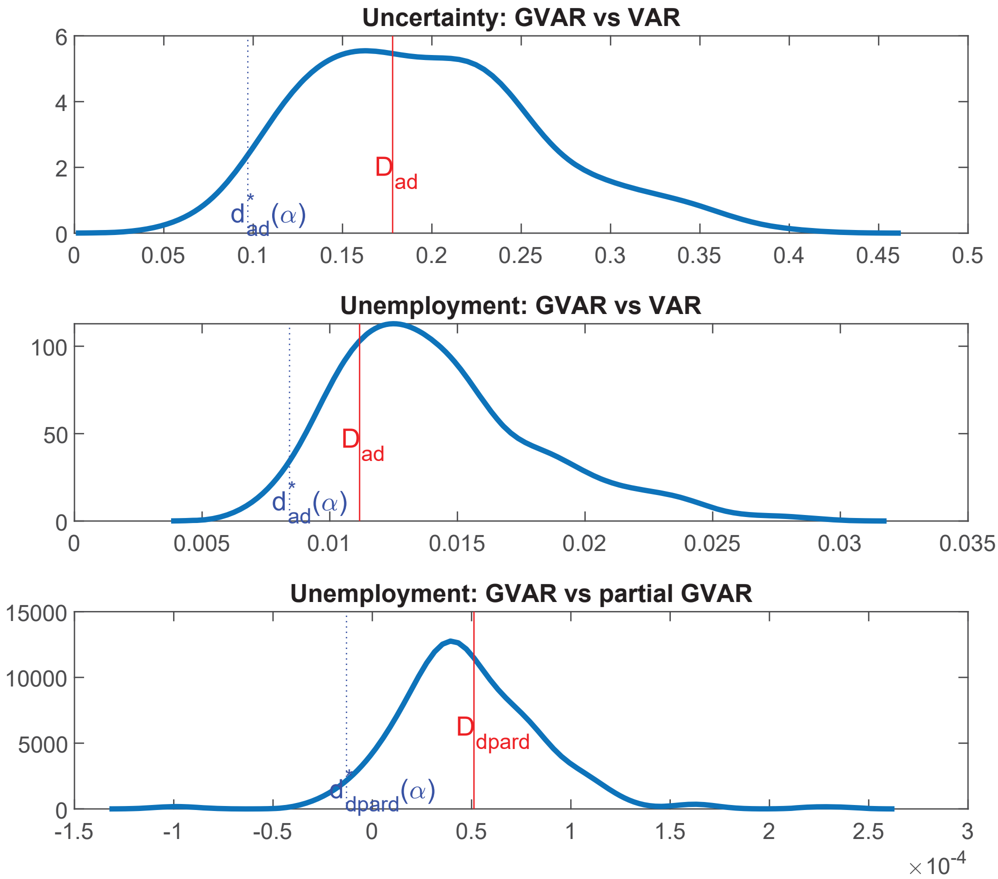


(a) Uncertainty



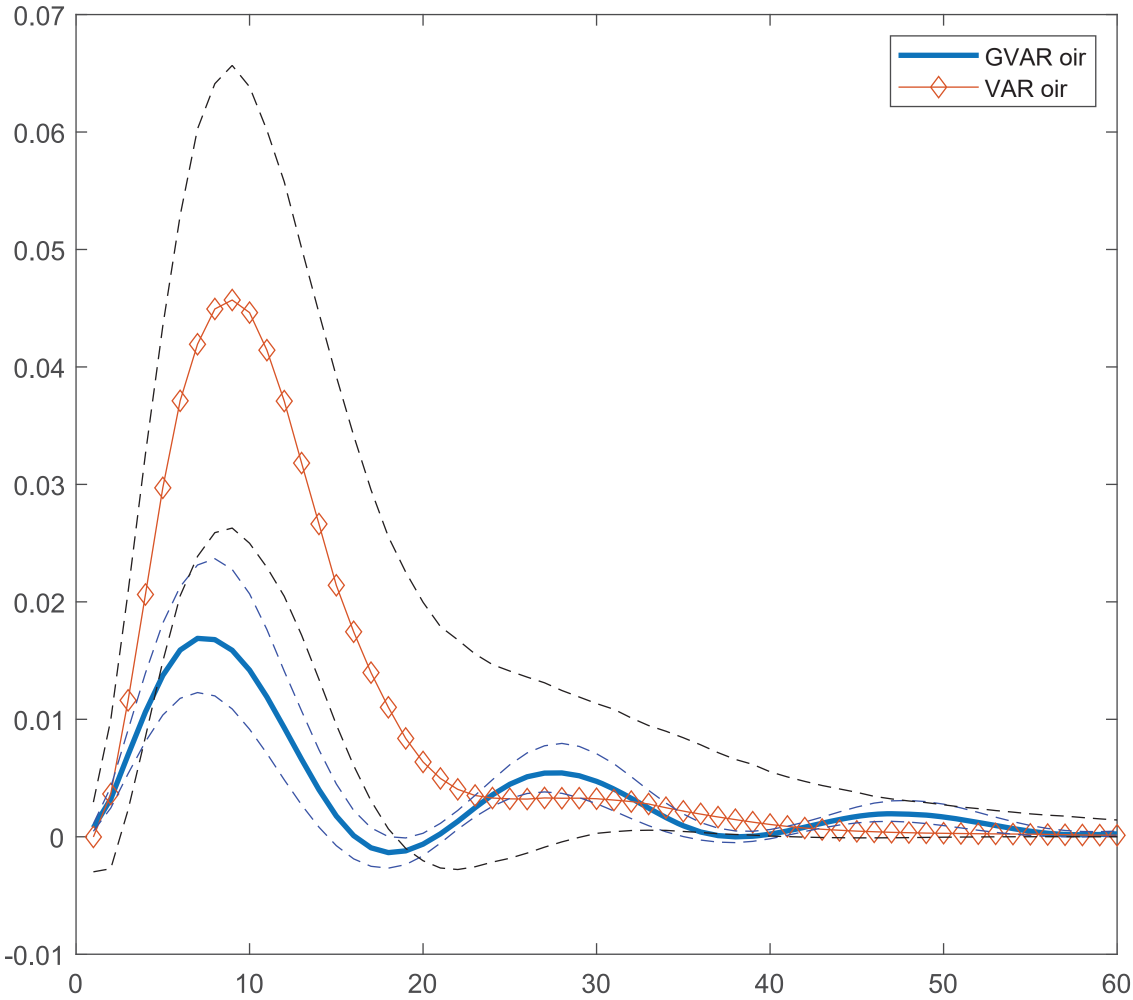
(b) Unemployment

Figure 2: Cross-state Uncertainty and Unemployment Correlations



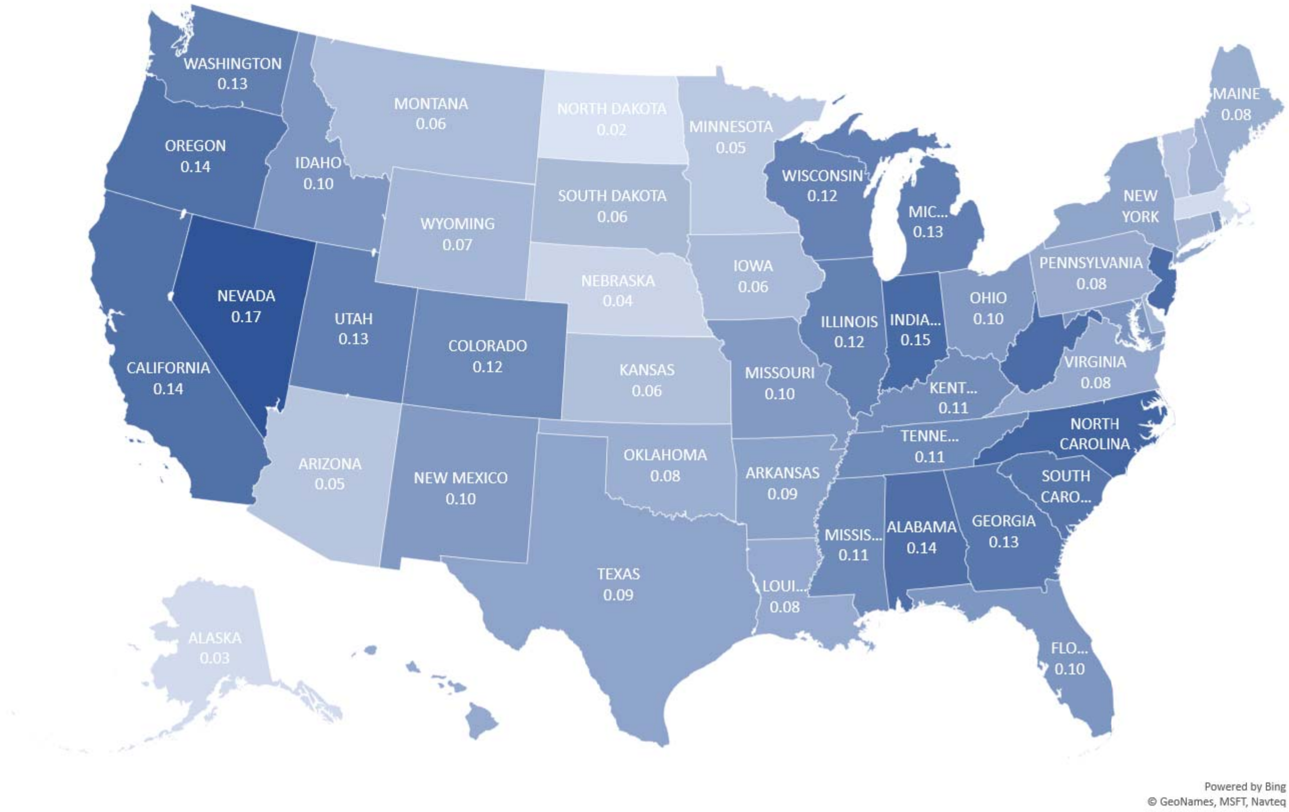
Notes: This figure plots the observed D_{ad}, which is the difference in the prediction criteria test statistic (red line) and $d^*_{ad}(\alpha)$, which is the left tail of the simulated distribution of the prediction criteria test statistic (dotted blue line) and assesses the ability of the disaggregate model versus the aggregate model in predicting aggregate uncertainty and aggregate unemployment respectively. The figure also plots the observed D_{dpard}, which is the difference in the prediction criteria test statistic (red line) and $d^*_{dpard}(\alpha)$, which is the left tail of the simulated distribution of the prediction criteria test statistic (dotted blue line) and assesses the ability of the disaggregate model versus the partial disaggregate model without state-level uncertainty in predicting unemployment.

Figure 3: Prediction Criteria



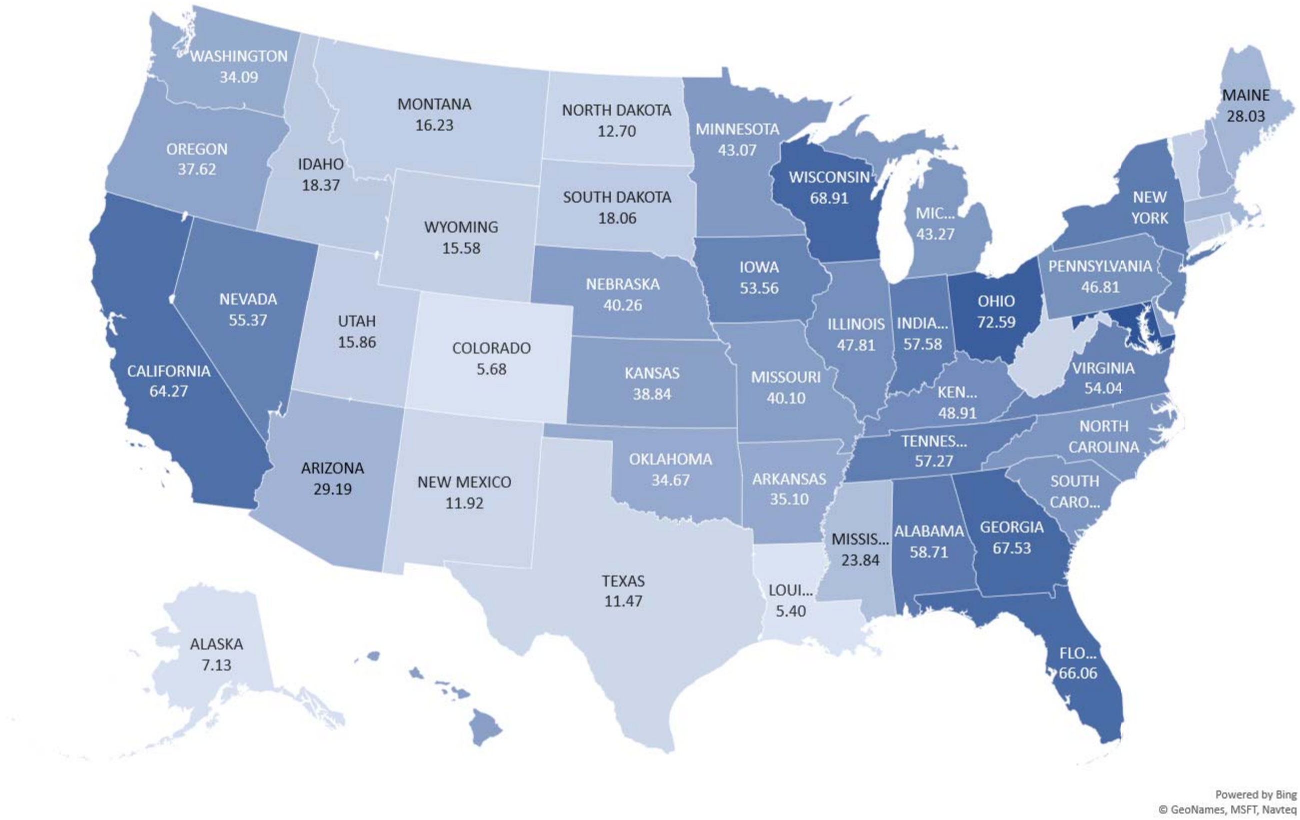
Notes: This figure compares the IRFs of an aggregate uncertainty shocks in the GVAR model to the VAR model. The GVAR IRFs are the weighted response of each state unemployment to an aggregate uncertainty shock in the GVAR model, where the aggregate shock is defined in Equation 17 through using population weights. The VAR IRFs are constructed according to Equation 8 via the Cholesky decomposition where uncertainty is placed first. Unemployment is defined as the quarterly change in unemployment rate. The confidence interval is the bootstrapped IRFs at ± 1 s.d.

Figure 4: Aggregate Uncertainty Shocks on Aggregate Unemployment



Notes: This figure presents the median estimate of the peak response of state-level unemployment to a one-standard-deviation aggregate uncertainty shock.

Figure 5: US Heatmap - Peak Response of State-Level Unemployment to an Aggregate Uncertainty Shock



Notes: This figure plots $\frac{p_{ii}^{NU}}{p_{ii}^{U}}$, which measures the relative importance of national influences in propagating uncertainty shocks on state i's unemployment.

Figure 6: US Heatmap - Relative Importance of National Uncertainty

Appendix of the paper "Uncertainty in a Disaggregate model: a Data Rich Approach using Google Search Queries" by Kalvinder Shields and Trung Duc Tran (Not for Publication)

The Appendix includes the following:

- A: Lists all sources of data used in this paper;
- B: Provides an example of how to construct a simple uncertainty index by using Google Trends;
- C: Presents the lists of words, which are associated with the divergence between uncertainty in California, Florida, New York, Texas and Virginia;
- D: Recovers the impulse responses of actual unemployment rate from the impulse responses of unemployment in quarterly difference;
- F: Robustness to a smaller set of states: namely excluding Delaware, (Washington) D.C, Vermont, Virginia;
- G: Alternative weighting approach: equal-weights, GDP-weights, distance weights;
- H: Robustness of cross-sectional regression on the heterogeneous effect of national uncertainty shocks.

A Data Sources

Data Sources for the Disaggregate Analysis:

- Unemployment: St Louis FRED.
- Population: For construction of weighting matrices. Source: Census Bureau.
- GDP: For construction of weighting matrices. Source: St Louis FRED.

Data sources for Cross-sectional Analysis:

- Industry shares of GDP: State-level GDP by industry, annual 2001 to 2016, average over time. Source: BEA. Industry classification is NAICS.

- Percentage of loans made by small bank: Bank balance sheet data on all FDIC-insured financial institutions excluding bank holding companies, 2001-2016. Small banks are defined as at or below the 90th percentile of the national distribution of bank size by assets. The small bank loans share is the time-average of the fraction of total loans on small bank balance sheets in each state. Source: Call Reports from the FFIEC.

- Homeownership rate: homeownership by state. 2001-2016, average over time: Source: St Louis FRED.

- Percentage of workers represented by unions: Union membership as a share of employment by state, average 2001-2016 Source: Hirsch and Macpherson (2003) (downloaded from unionstats.com).

- Right-To-Work (RTW): Dummy for whether a state has right to work legislation as of 2016. Source: http://www.nrtw.org/right-to-work-states.

- State-level debt-GDP ratio: average 2001-2016. Source: Census Bureau.

- State-level government spending-GDP ratio: average 2001-2016. Source: Census Bureau.

- State-level government tax revenue-GDP ratio: average 2001-2016. Source: Census Bureau.

- Intergovernmental transfer: federal spending transferred to state and local governments. Source: Census Bureau.

- Share of internet users: percent of the population used the internet by state in 2017. Source: Statista.

B Constructing the Google Trends Uncertainty Index

This section provides an example that helps to clarify the general representation of the Google Trends uncertainty index construction. Let's assume we only need three search terms to construct the uncertainty index: *banking*, *default* and *presidential election*.

The first three columns in A1 show the search output returned by Google Trends when each of these terms is searched separately. As mentioned earlier, the resulting outputs by Google Trends are scaled to a range between 0 and 100 when the search frequency is highest.

Things may change when two search terms are searched together. In this case, *banking* is chosen to be the benchmark term. For example, when *banking* is searched together with *default*, the search frequency for *default* (column 5) is not the same than that when *default* is search separately (column 2) because the resulting search frequency is scaled when the search frequency for *banking* is the highest in April 2010. Similarly, the search term for *banking* (column 4) when searched together with *presidential election* is not the same when it is searched separately (column 1).

As a result, by using the search frequency for *banking* as a benchmark when it is searched with *default*, the search frequency for *presidential election* can then be scaled accordingly. Column 8-10 present the final standardised search frequency that is ready to aggregate to construct the uncertainty index that are made of three search terms.

C Search Terms Associated with the Difference between State Uncertainty and National Uncertainty

In Figure 1 we can see that there are several periods where state uncertainty and national uncertainty do not comove. Table A2 selects some highlighted periods and the keywords are associated with the divergence between state uncertainty and national uncertainty in each period. We can see that most of differences are related to the search terms are about the budget, the financial criss, the price level, the job market and reform.

D Recovering Impulse Responses for the Actual Unemployment Rate

When unemployment is specified as the quarterly difference, when we plot it in non-cumulated terms, we have:

$$IRF_t^Q = IRF_t - IRF_{t-2}$$

where IRF_t^Q is the response of the quarterly-differenced unemployment rate, and IRF_t is the response of the actual unemployment rate. As such, we can recover the response of the actual unemployment rate as follows:

$$IRF_1^Q = IRF_1 - IRF_{-1} = IRF_1(*)$$

$$IRF_2^Q = IRF_2 - IRF_0 = IRF_2$$

This is because of the initial condition is the same as in the case in which there is no shock.

$$IRF_3^Q = IRF_3 - IRF_1$$

From $(*)$, we can then recover IRF_3. By following the same logic, we can then recover the responses of the actual unemployment rate across all different horizons.

E Additional Results

Figure A1 reports the ratio: $\frac{p_{ii}^{NU}}{p_{ii}^{U}}$, (with estimation errors) which gives the relative importance of the above-defined uncertainty shocks propogating through national dynamics relative to propogating through both state and national dynamics, as modelled in Equation 7, with respect to the persistent effect on the unemployment variable in each state, scaled by a standard error state-specific unemployment shock on impact. On average, the uncertainty shocks working their way through the channels associated with national elements account for 37% of the total variation in state-specific unemployment and the remaining 53% is attributed to the purely state-specific channel abstracting entirely from national dynamics.

F Subset of States

Due to the lack of comovement in Delaware, D.C, Vermont and Virginia, we remove these states from the analysis. We find that the main results are still unchanged. Figure A2 shows that the disaggregate model can predict the data better than the aggregate model. Figure A3 shows that the effect of aggregate uncertainty shocks on aggregate unemployment captured by the disaggregate model is still smaller than the aggregate model, yet the effect is relatively larger than the baseline model.Figure A4 shows that the average impact of national uncertainty shocks are still less than state-specific uncertainty shocks.

G Alternative Weighting Schemes

In this section, we provide additional robustness checks on how the results might change if we use a different method to construct the aggregate variable and link individual states together as a system.

Distance weights: In the first check, we use distance weight to link state together. The distant weights are constructed in two steps. First, we measure the distance between each state capital cities and inverse these numbers. Second, we rescale to ensure the weights add up to 1. The matrix plays an important role in providing meaningful links between states and allowing a measure of the degree of dependency between states. For example, New Jersey, Connecticut and Pennsylvania have much larger weight on New York than Arizona or Texas. Based on this matrix, we can then construct the *starred* variables for unemployment rate and uncertainty. The main results are in Figure A5, A6 and A7.

Equal weights: We employ an equal weighting approach to construct *starred* variables and link states. The results can be found in Figure A8, A9 and A10. We find that first the GVAR IRF is larger than the baseline model but is still less than the VAR IRF. Second, the relative importance of national uncertainty shocks on unemployment is larger than the baseline approach but on average, national uncertainty is still less important than state-specific uncertainty.

GDP weights: We employ a GDP weighting approach to construct *starred* variables and link states. The results can be found in Figure A11, A12 and A13. We find that the results are almost identical to the baseline approach.

H Cross-Sectional Regressions

Table A3 and A4 present the regression results on the selection of the industry variables to be included in the baseline model. We sequentially include one industry at a time. We retain if the industry variable is statistically significant at the 10% level of significance. Using this strategy, we find that only the manufacturing industry and the mining industry are statistically insignificant in their association with the state-level response of unemployment to uncertainty shock, as presented in A3 Column 6. On the other hand, we find that only the government industry and the real estate industry have a statistically significant relationship in the relative importance of national influences in propagating the uncertainy shocks, as presented in A4 Column 8.

Table A1: Constructing the GTU Index: An Illustration

				Searched together		Searched together		Final		
Time	banking [1]	default [2]	presidential election [3]	banking [4]	default [5]	banking [6]	presidential election [7]	banking [8]	default [9]	presidential election [10]
2004-01	28	60	4	28	21	24	4	28	21	4.67
2004-02	27	64	5	27	22	24	5	27	22	5.63
2004-03	23	60	4	23	26	20	4	23	26	4.60
2008-09	45	100	8	45	31	39	8	45	31	9.23
2008-10	41	99	14	41	31	35	14	41	31	16.40
2008-11	43	91	22	43	27	37	22	43	27	25.57
2010-04	100	76	1	100	25	86	1	100	25	1.16
2010-05	95	69	1	95	23	81	1	95	23	1.17
2016-11	40	47	100	40	15	35	100	40	15	114.29
2016-12	38	46	4	38	15	33	4	38	15	4.61
2017-01	40	50	3	40	17	34	3	40	17	3.53
2018-04	39	47	1	39	16	33	1	39	16	1.18
2018-05	38	49	1	38	15	33	1	38	15	1.15

The first three columns show the relative search frequency provided by Google Trends when the search term *banking*, *default*, and *presidential election* are searched separately. The next two columns show the search frequency provided by Google Trends when the search term *banking* and *default* are searched together. The next two columns show the search frequency provided by Google Trends when the search term *banking* and *presidential election* are searched together. The last three columns show the search frequency provided by Google Trends after the transformation described in the paper.

Table A2: Search Terms Associated with the Difference between State Uncertainty and National Uncertainty

California					
Time	Keywords				
Jan-06	Millitary budget	Real estate bubble	Trade policy	Pollution control	Federal Fund rate
Aug-06	Food price	Real estate bubble	Job security	Pollution control	Food and drug aministration
Sep-10	Financial crisis	Share price	Food price	austerity	Monetary Policy
Nov-10	Financial crisis	Tax cut	Unemployment benefit	Recession	Fiscal policy
Dec-10	Financial crisis	Healthcare	Pollution control	Unemployment benefit	United States Congress
Florida					
Time	Keywords				
Feb-06	Tax law	Tax cut	Worker's compensation	Home price	Stock exchange
Apr-08	Fuel price	Recession	Job Security	Foreclosure	Presidential election
May-09	Food price	Monetary policy	Recession	Share price	Job security
Jun-13	Banking	Foreclosure	Job Security	Share price	White House
Sep-17	Banking	Bank Loan	Share price	Home price	Job security
New York					
Time	Keywords				
Jun-11	austerity	debt ceiling	Financial crisis	healthcare reform	unemployment benefit
Oct-12	Price fuel	Unemployment benefit	Fuel price	healthcare reform	Food price
Aug-16	Food price	Job security	Presidential election	Share Price	Financial crisis
Jun-17	Share price	Job security	Gas price	Fuel price	food price
Nov-17	Tax cut	Fuel price	National debt of United state	Fiscal policy	Financial crisis
Texas					
Time	Keywords				
Oct-06	Monetary policy	Fiscal policy	Recession	Presidential election	Food and Drug Administration
Feb-07	Food price	Share price	Tax cut	Energy policy	Business outlook
Sep-07	Job security	Share price	recession	Presidental election	Bank rate
Jul-09	Financial crisis	Healthcare reform	Recession	Security and exchange commission	Money supply
May-12	Foreclosure	Healthcare reform	Healthcare act	Fuel price	austerity
Virginia					
Time	Keywords				
Jul-06	Equal opportunity employment	Fuel price	Job security	Bank of England	Food price
Nov-10	Financial Crisis	Healthcare reform	Monetary policy	Share price	Banking
Sep-12	Austerity	Financial Crisis	Food price	Job security	Millitary budget
Mar-15	Share price	Financial Crisis	Fuel price	Job security	Money Supply
Aug-16	Healthcare act	Financial Crisis	Fuel price	Share price	Millitary budget

Table A3: Heterogeneity Analysis of the Response of State-Level Unemployment to an Aggregate Uncertainty Shock: Industry Mix

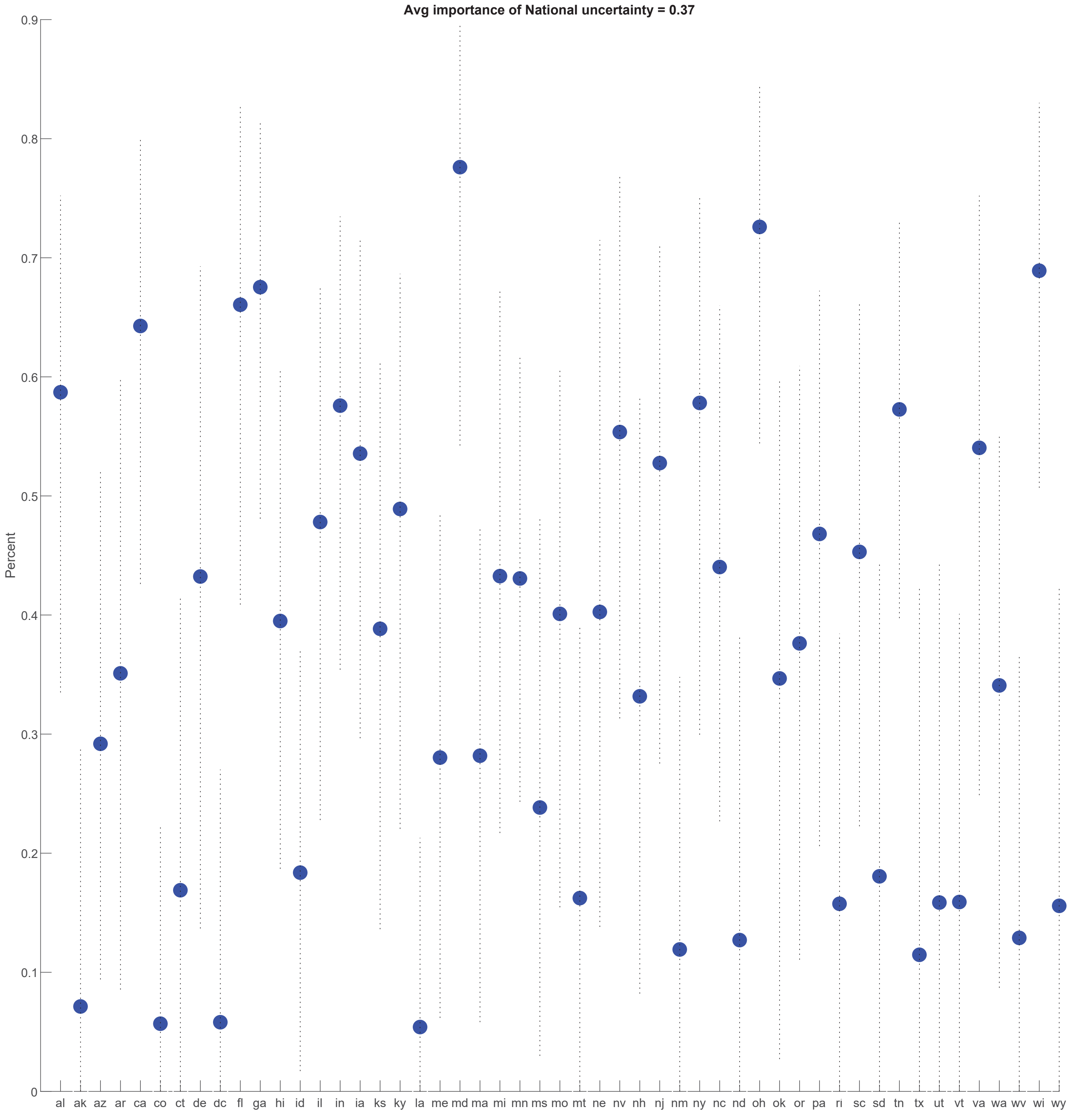
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
VARIABLES	response	response	response	response	response	response	response
Agriculture	-0.00502						
	(0.00343)						
Construction		0.00132					
		(0.00620)					
Finance			0.000234				
			(0.00108)				
Government				-0.00130			
				(0.000908)			
Manufacturing					0.00204**	0.00152**	0.00172*
					(0.000856)	(0.000749)	(0.000925)
Mining						-0.00164*	-0.00135
						(0.000905)	(0.000955)
RealEstate							0.00114
							(0.00231)
Observations	51	51	51	51	51	51	51
Adjusted R-squared	0.360	0.335	0.335	0.359	0.416	0.469	0.459

Notes: $response_i$ is defined in Section 5.3 and represents the peak state-level response of unemployment to an aggregate uncertainty shock. Robust standard errors in parentheses $^{***}p < 0.01, ^{**}p < 0.05, ^{*}p < 0.1$.

Table A4: Heterogeneity Analysis of the Importance of National Influences in Propagating Uncertainty Shocks: Industry Mix

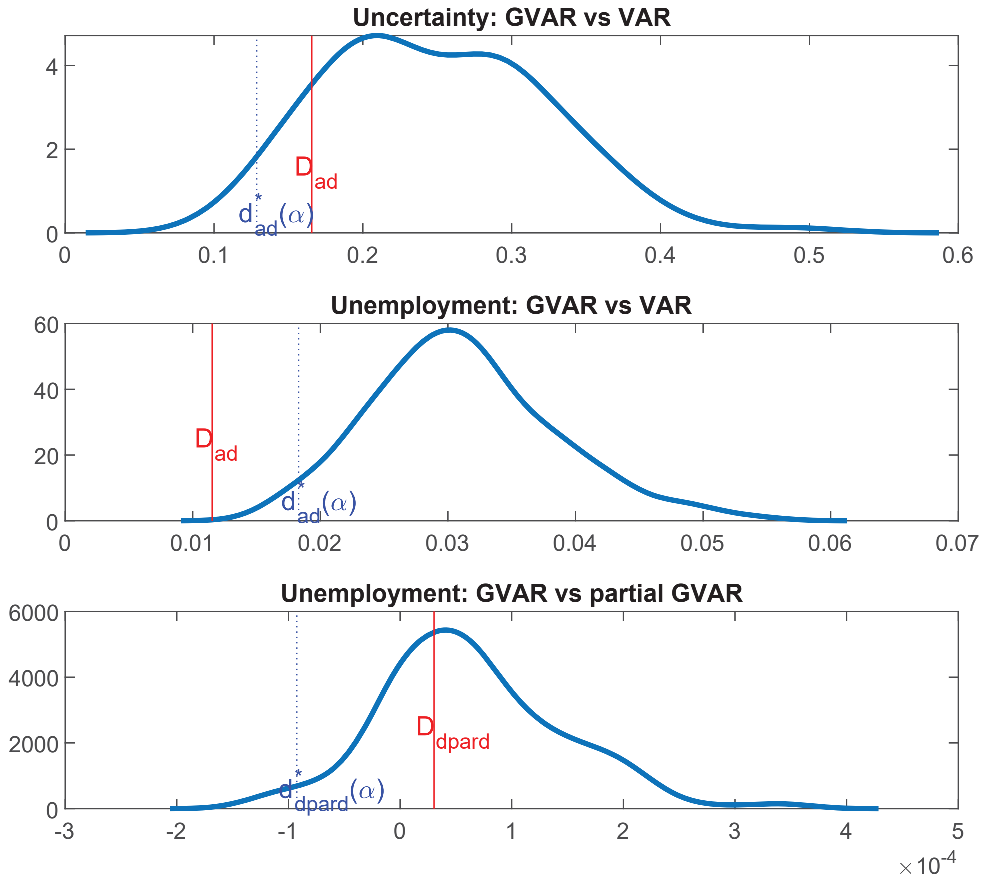
VARIABLES	(1) $\frac{p_{ii}^{NU}}{p_{ii}^{U}}$	(2) $\frac{p_{ii}^{NU}}{p_{ii}^{U}}$	(3) $\frac{p_{ii}^{NU}}{p_{ii}^{U}}$	(4) $\frac{p_{ii}^{NU}}{p_{ii}^{U}}$	(5) $\frac{p_{ii}^{NU}}{p_{ii}^{U}}$	(6) $\frac{p_{ii}^{NU}}{p_{ii}^{U}}$	(7) $\frac{p_{ii}^{NU}}{p_{ii}^{U}}$	(8) $\frac{p_{ii}^{NU}}{p_{ii}^{U}}$
Agriculture	-2.090							
	(1.418)							
Construction		3.521						
		(4.647)						
Finance			0.661					
			(0.853)					
Government				-1.406**	-1.503**	-1.318*	-1.041**	-1.016**
				(0.638)	(0.665)	(0.654)	(0.446)	(0.420)
Manufacturing					-0.191			
					(0.532)			
Mining						-1.028*	-0.391	
						(0.551)	(0.563)	
RealEstate							3.112***	3.655***
							(1.108)	(0.947)
Observations	51	51	51	51	51	51	51	51
Adjusted R-squared	0.277	0.283	0.278	0.354	0.340	0.429	0.505	0.508

Notes: $\frac{p_{ii}^{NU}}{p_{ii}^{U}}$ is defined in Section 5.4 and represents the relative importance of national uncertainty shocks in propagating uncertainty shocks. Robust standard errors in parentheses $^{***}p < 0.01,^{**}p < 0.05,^{*}p < 0.1$.



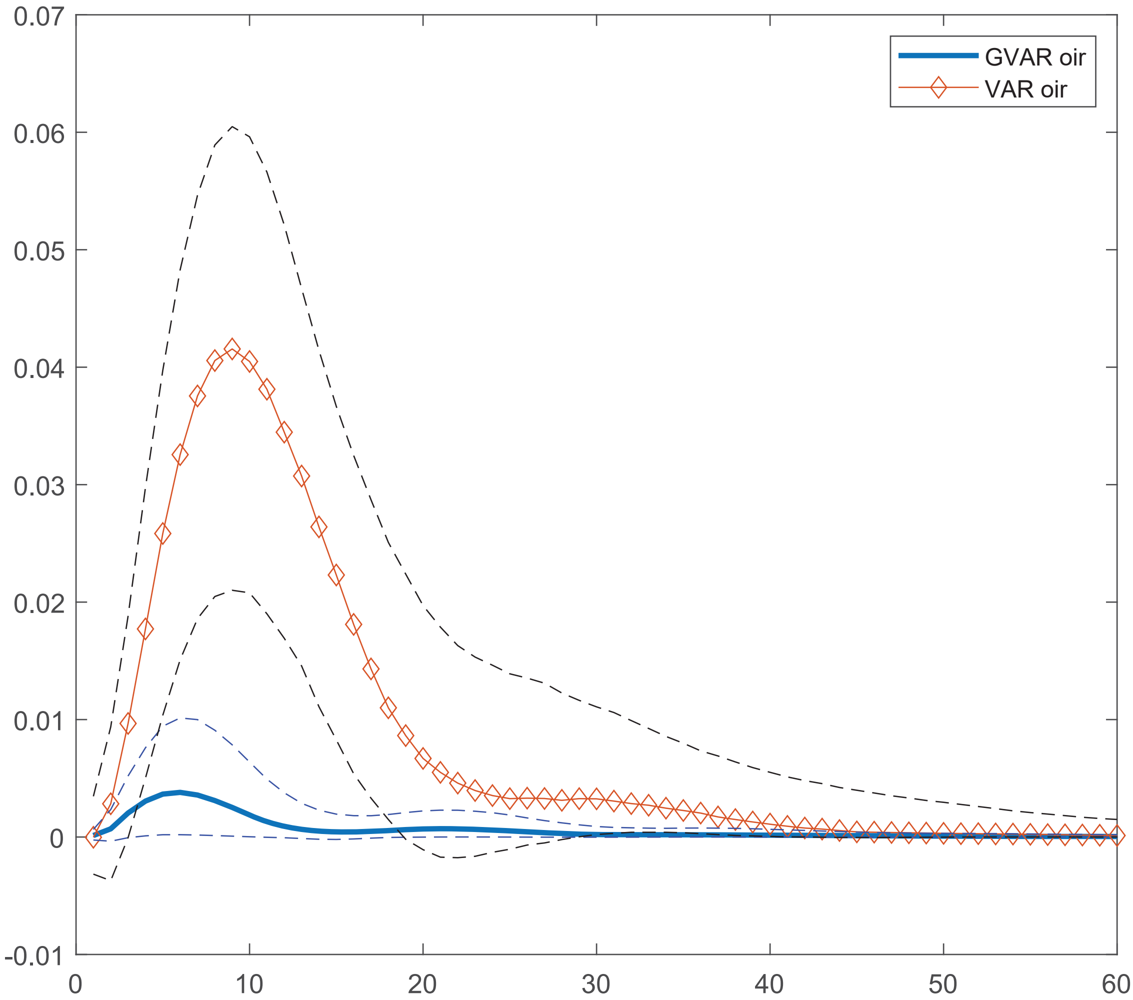
Notes: This figure plots $\frac{p_{ii}^{NU}}{p_{ii}^{U}}$, which measures the relative importance of national influences in propagating uncertainty shocks on state i's unemployment (the blue dot). The 95 % bootstrapped confidence interval are denoted by the dotted lines.

Figure A1: Relative Importance of National Uncertainty



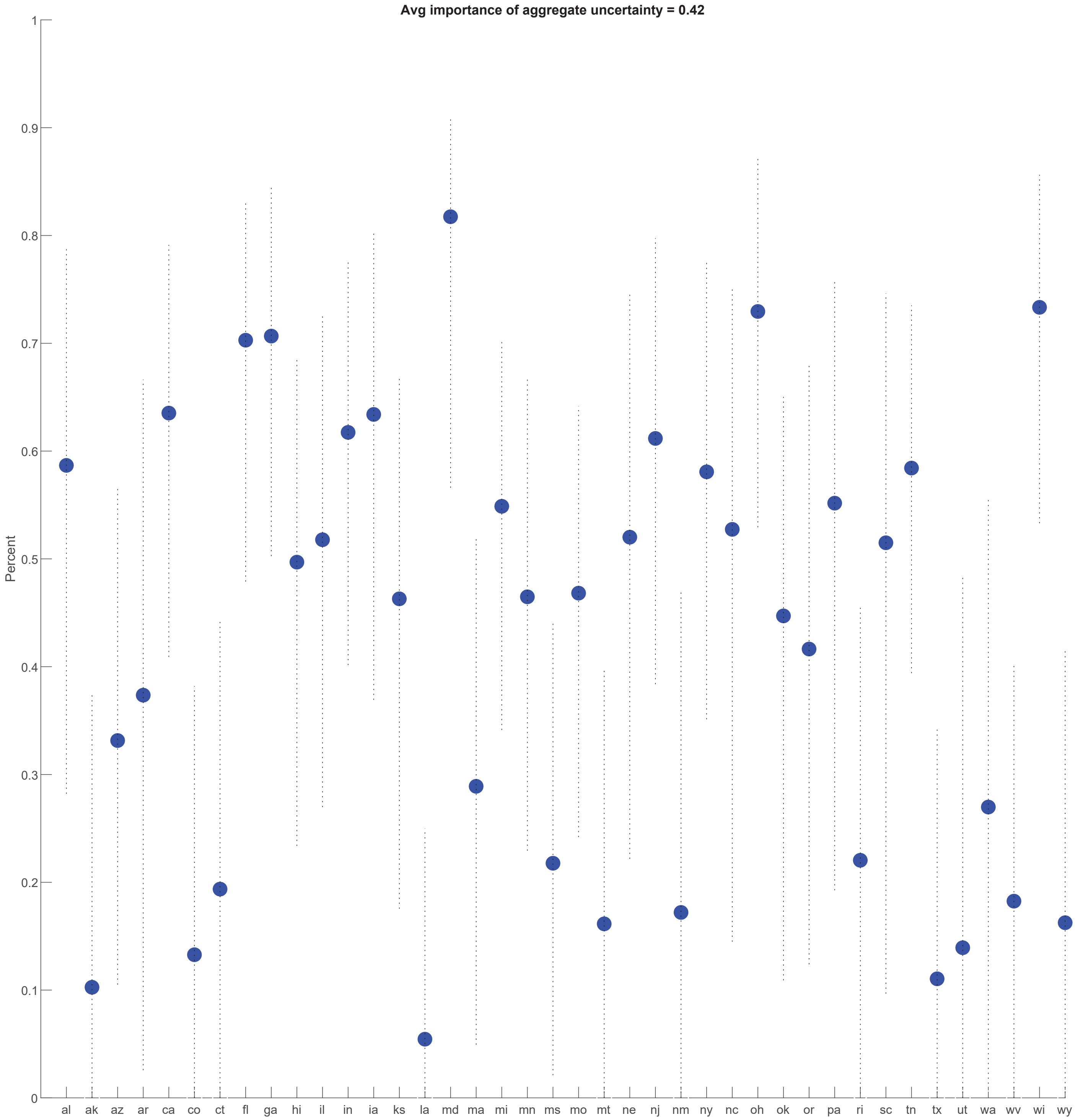
Notes: This figure plots the observed D_{ad}, which is the difference in the prediction criteria test statistic (red line) and $d^*_{ad}(\alpha)$, which is the left tail of the simulated distribution of the prediction criteria test statistic (dotted blue line) and assesses the ability of the disaggregate model versus the aggregate model in predicting aggregate uncertainty and aggregate unemployment respectively. The figure also plots the observed D_{dpard}, which is the difference in the prediction criteria test statistic (red line) and $d^*_{dpard}(\alpha)$, which is the left tail of the simulated distribution of the prediction criteria test statistic (dotted blue line) and assesses the ability of the disaggregate model versus the partial disaggregate model without state-level uncertainty in predicting unemployment.

Figure A2: Prediction Criteria - Robustness Analysis using Distance Weights



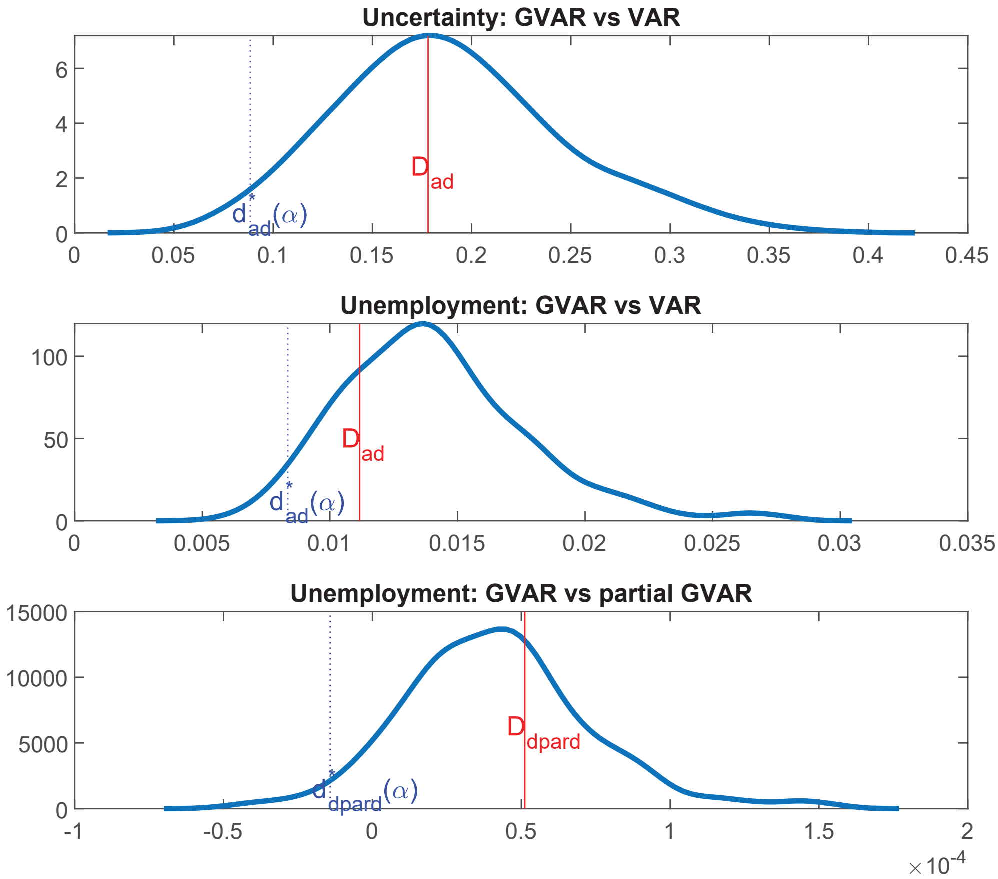
Note: This figure compares the IRFs of an aggregate uncertainty shocks in the GVAR model to the VAR model. The GVAR IRFs are the weighted response of each state unemployment to an aggregate uncertainty shock in the GVAR model, where the aggregate shock is defined in Equation 17 through using population weights. The VAR IRFs are constructed according to Equation 8 via the Cholesky decomposition where uncertainty is placed first. Unemployment is defined as the quarterly change in unemployment rate. The confidence interval is the bootstrapped IRFs at ±1 s.d.

Figure A3: Aggregate Uncertainty Shocks on Aggregate Unemployment - Robustness Analysis using Fewer States



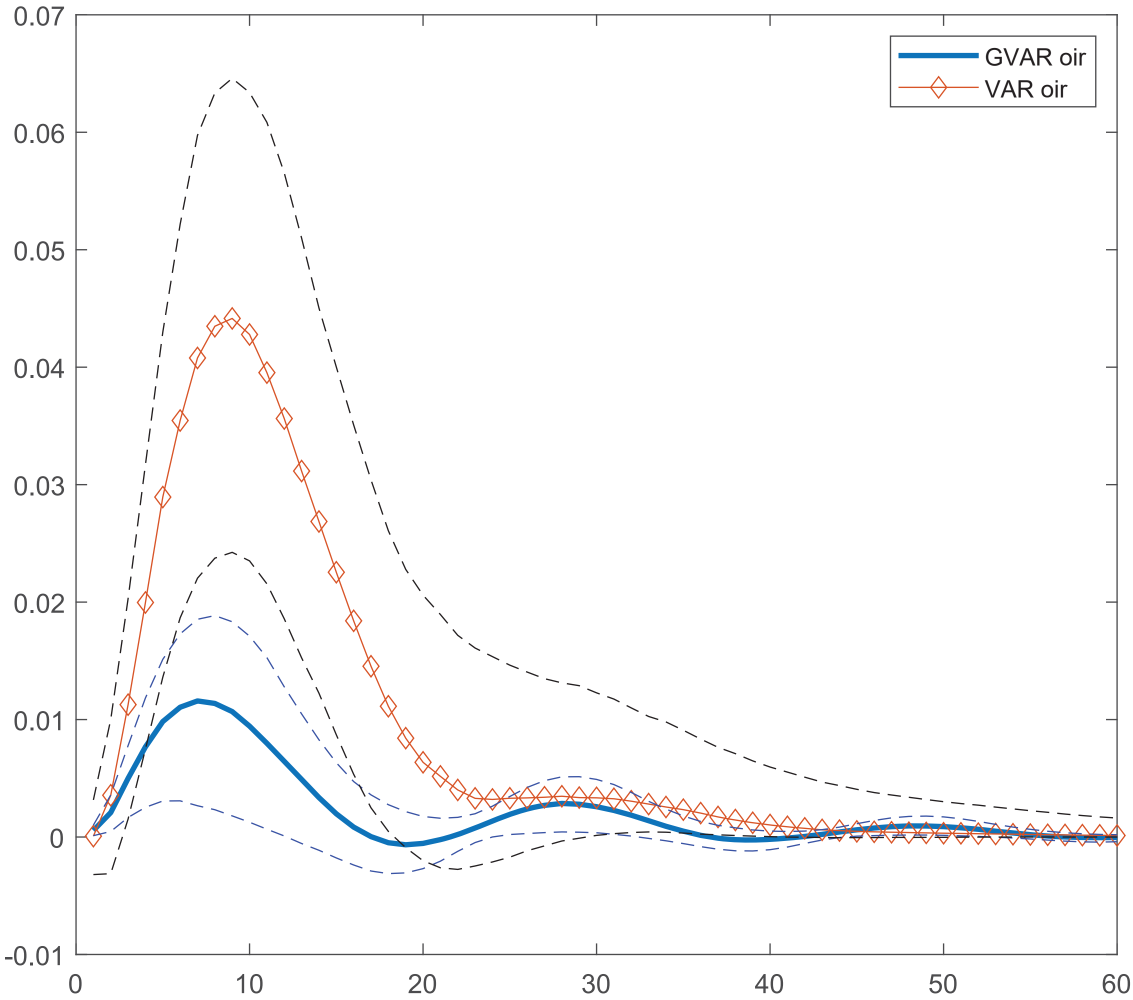
Notes: This figure plots $\frac{p_{ii}^{NU}}{p_{ii}^{U}}$, which measures the relative importance of national influences in propagating uncertainty shocks on state i's unemployment (the blue dot). The 95 % bootstrapped confidence interval are denoted by the dotted lines.

Figure A4: Relative Importance of National Uncertainty - Robustness Analysis using Fewer States



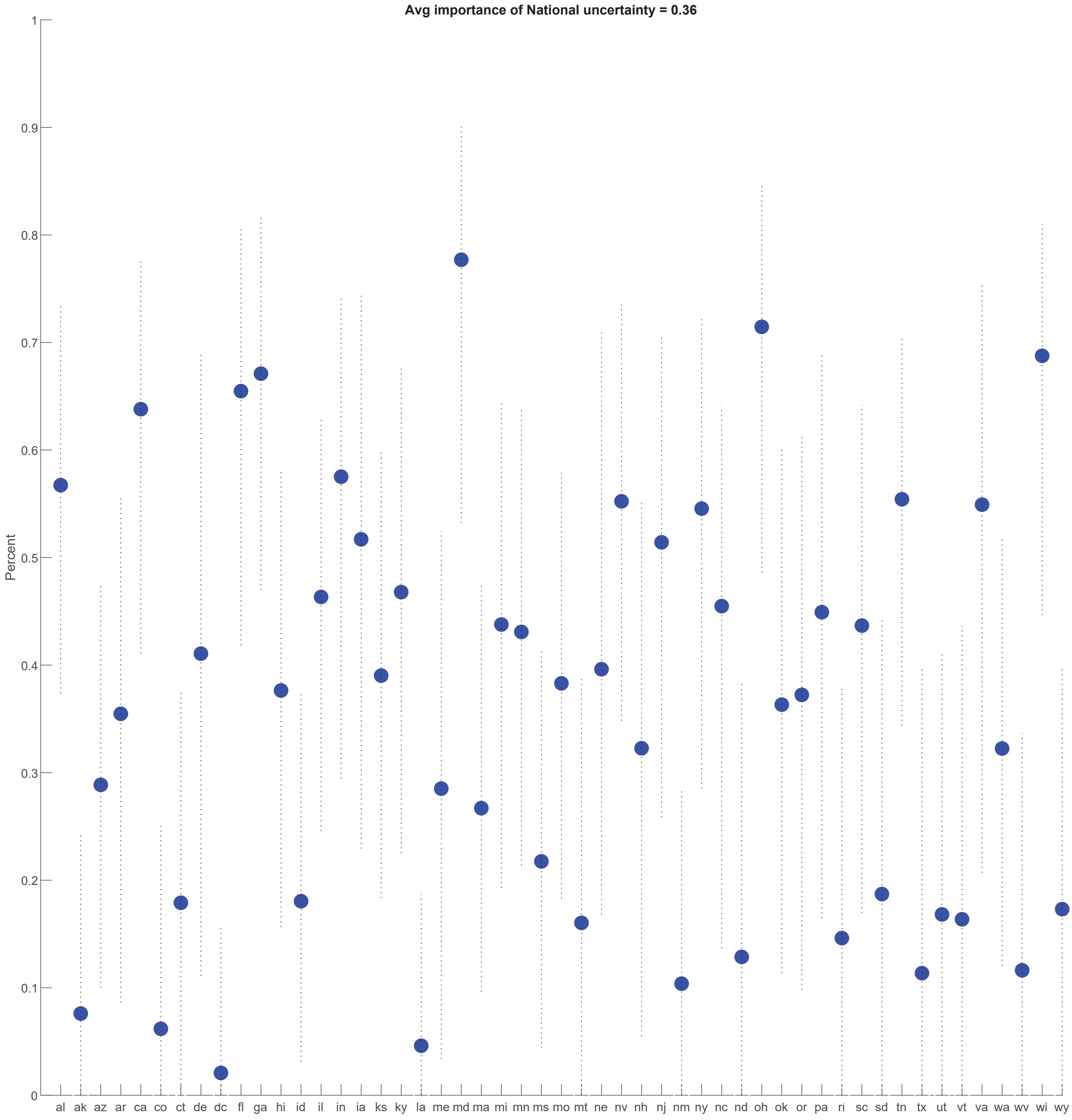
Notes: This figure plots the observed D_{ad}, which is the difference in the prediction criteria test statistic (red line) and $d^*_{ad}(\alpha)$, which is the left tail of the simulated distribution of the prediction criteria test statistic (dotted blue line) and assesses the ability of the disaggregate model versus the aggregate model in predicting aggregate uncertainty and aggregate unemployment respectively. The figure also plots the observed D_{dpard}, which is the difference in the prediction criteria test statistic (red line) and $d^*_{dpard}(\alpha)$, which is the left tail of the simulated distribution of the prediction criteria test statistic (dotted blue line) and assesses the ability of the disaggregate model versus the partial disaggregate model without state-level uncertainty in predicting unemployment.

Figure A5: Prediction Criteria - Robustness Analysis using Distance Weights



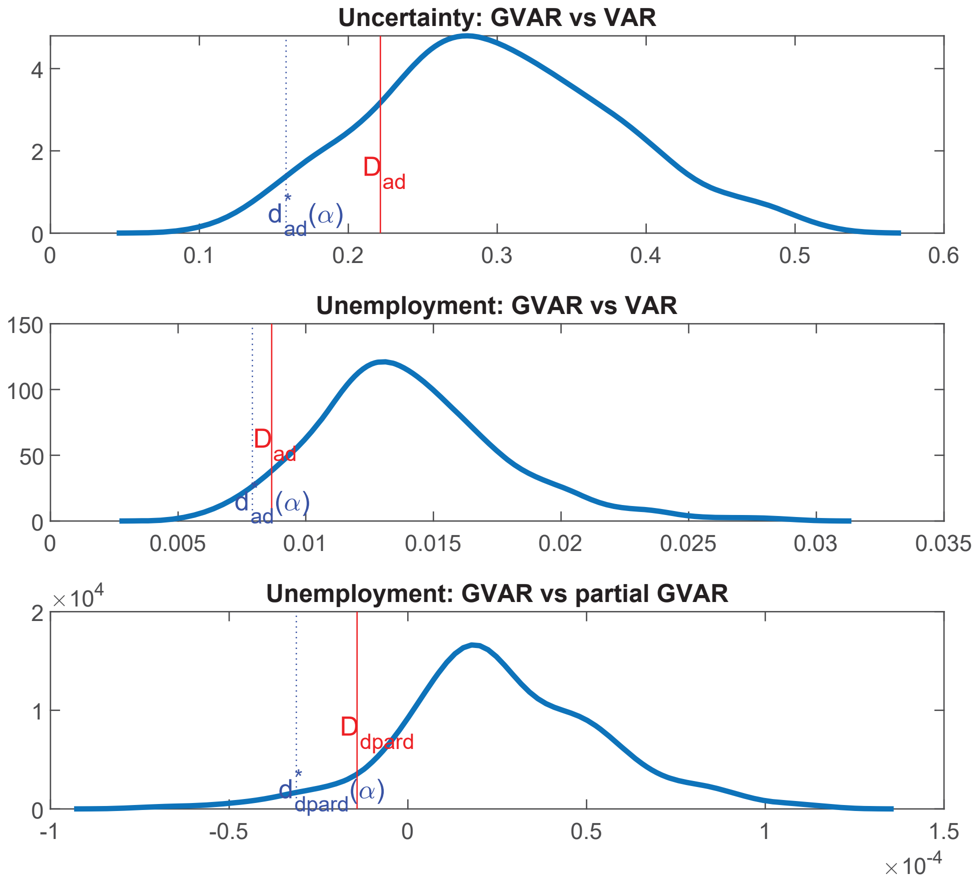
Note: This figure compares the IRFs of an aggregate uncertainty shocks in the GVAR model to the VAR model. The GVAR IRFs are the weighted response of each state unemployment to an aggregate uncertainty shock in the GVAR model, where the aggregate shock is defined in Equation 17 through using population weights. The VAR IRFs are constructed according to Equation 8 via the Cholesky decomposition where uncertainty is placed first. Unemployment is defined as the quarterly change in unemployment rate. The confidence interval is the bootstrapped IRFs at ± 1 s.d.

Figure A6: Aggregate Uncertainty Shocks on Aggregate Unemployment - Robustness Analysis using Distance Weights



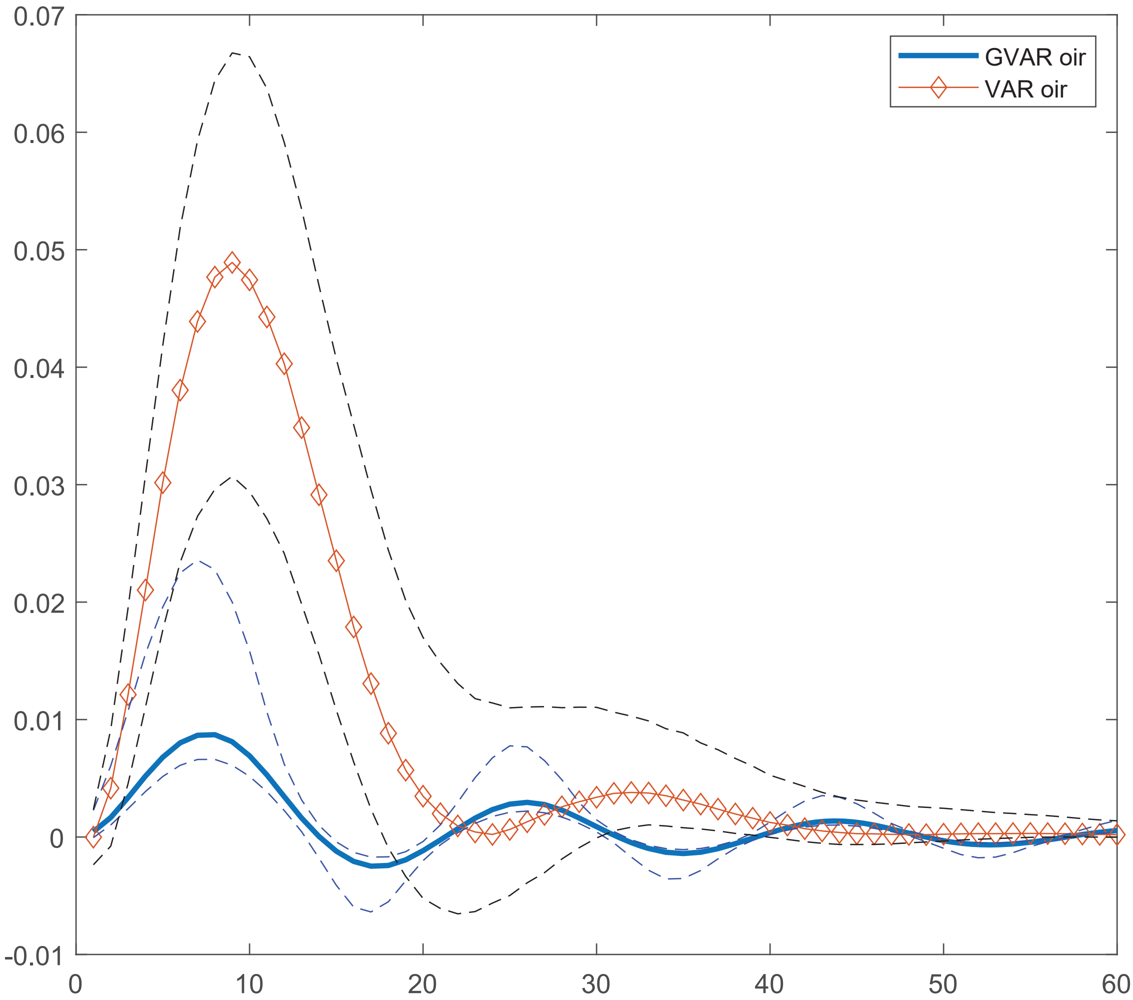
Notes: This figure plots $\frac{p_{ii}^{NU}}{p_{ii}^{U}}$, which measures the relative importance of national influences in propagating uncertainty shocks on state i's unemployment (the blue dot). The 95 % bootstrapped confidence interval are denoted by the dotted lines.

Figure A7: Relative Importance of National Uncertainty - Robustness Analysis using Distance Weights



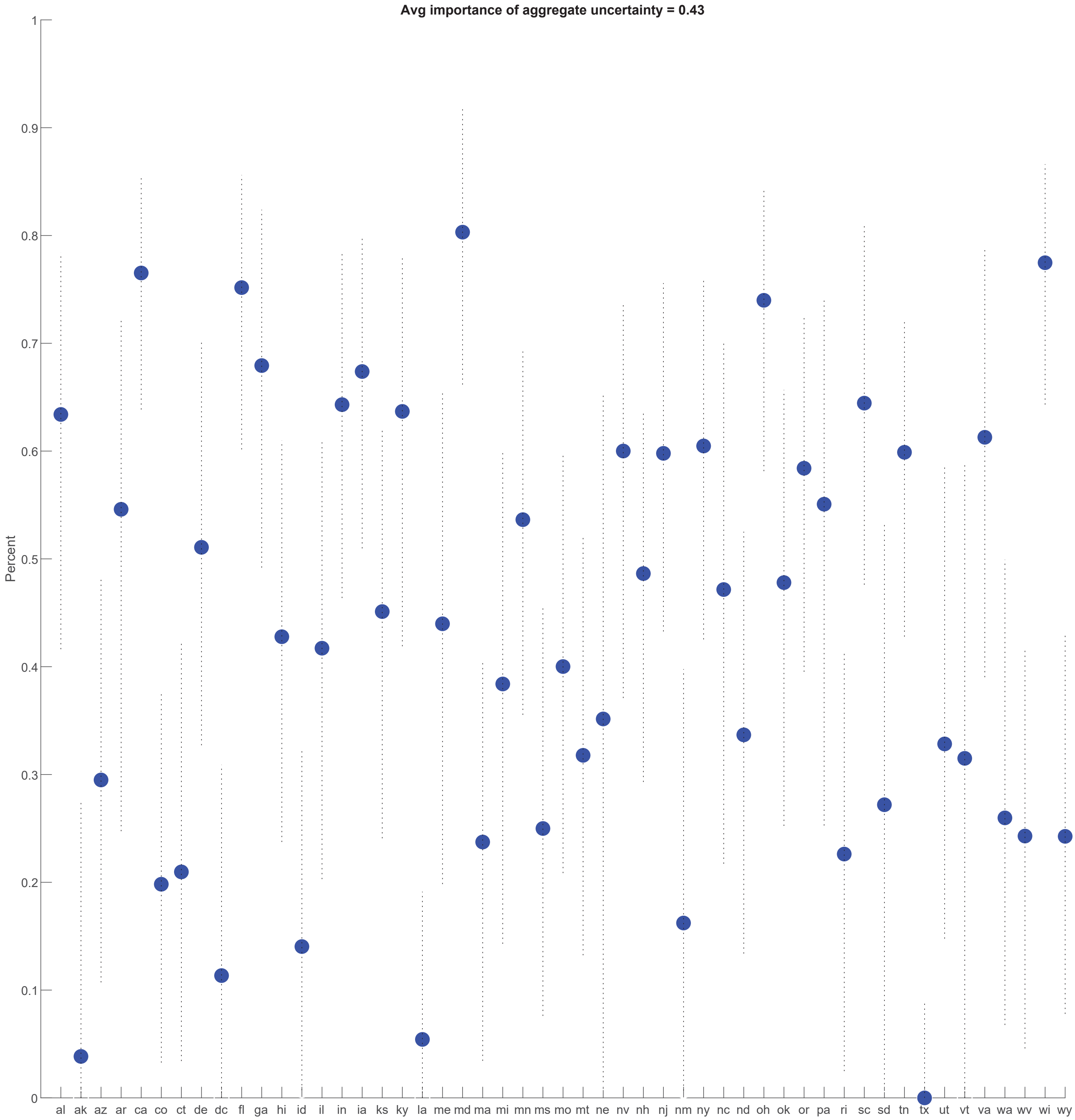
Notes: This figure plots the observed D_{ad}, which is the difference in the prediction criteria test statistic (red line) and $d^*_{ad}(\alpha)$, which is the left tail of the simulated distribution of the prediction criteria test statistic (dotted blue line) and assesses the ability of the disaggregate model versus the aggregate model in predicting aggregate uncertainty and aggregate unemployment respectively. The figure also plots the observed D_{dpard}, which is the difference in the prediction criteria test statistic (red line) and $d^*_{dpard}(\alpha)$, which is the left tail of the simulated distribution of the prediction criteria test statistic (dotted blue line) and assesses the ability of the disaggregate model versus the partial disaggregate model without state-level uncertainty in predicting unemployment.

Figure A8: Prediction Criteria - Robustness Analysis using Equal Weights



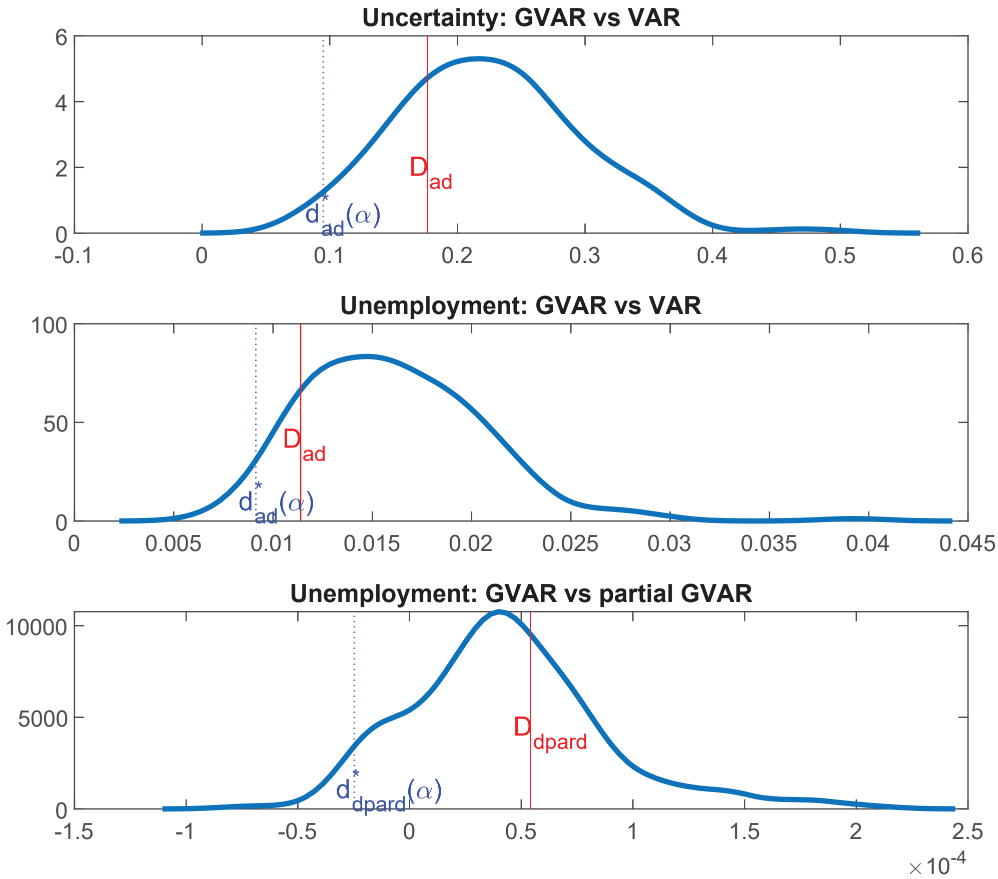
Note: This figure compares the IRFs of an aggregate uncertainty shocks in the GVAR model to the VAR model. The GVAR IRFs are the weighted response of each state unemployment to an aggregate uncertainty shock in the GVAR model, where the aggregate shock is defined in Equation 17 through using population weights. The VAR IRFs are constructed according to Equation 8 via the Cholesky decomposition where uncertainty is placed first. Unemployment is defined as the quarterly change in unemployment rate. The confidence interval is the bootstrapped IRFs at ± 1 s.d.

Figure A9: Aggregate Uncertainty Shocks on Aggregate Unemployment - Robustness Analysis using Equal Weights



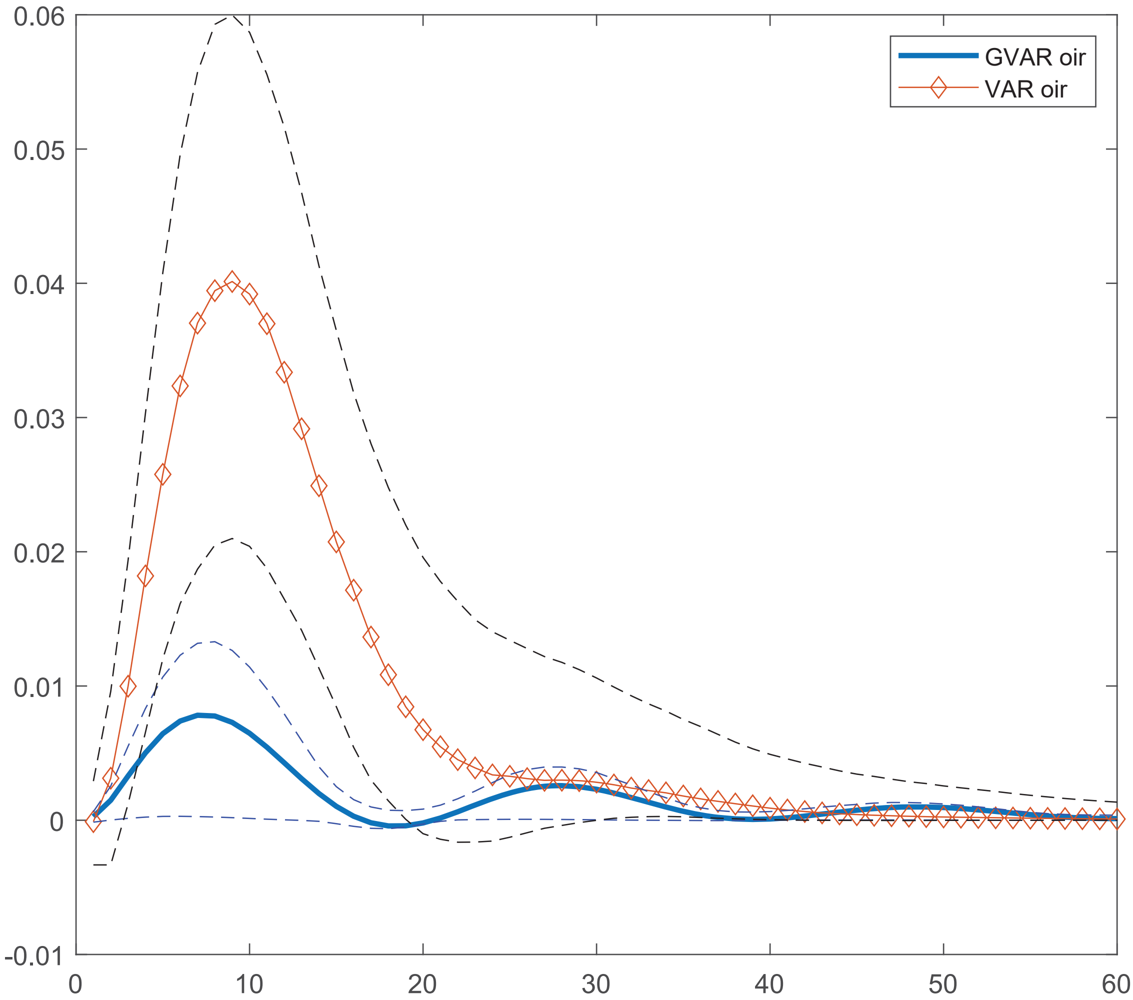
Notes: This figure plots $\frac{p_{ii}^{NU}}{p_{ii}^{U}}$, which measures the relative importance of national influences in propagating uncertainty shocks on state i's unemployment (the blue dot). The 95 % bootstrapped confidence interval are denoted by the dotted lines.

Figure A10: Relative Importance of National Uncertainty versus State-Level Shocks - Robustness Analysis using Equal Weights



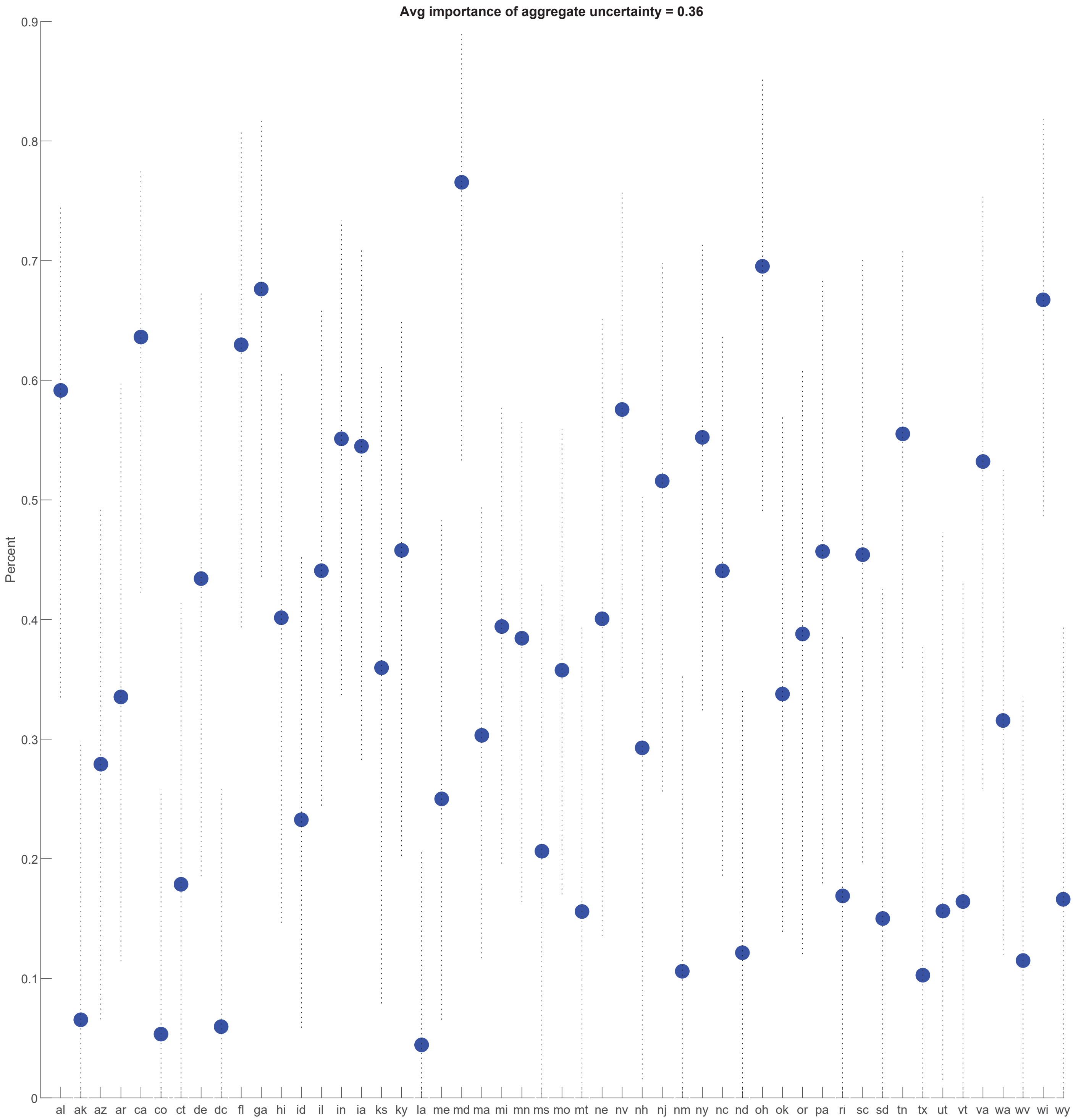
Notes: This figure plots the observed D_{ad}, which is the difference in the prediction criteria test statistic (red line) and $d^*_{ad}(\alpha)$, which is the left tail of the simulated distribution of the prediction criteria test statistic (dotted blue line) and assesses the ability of the disaggregate model versus the aggregate model in predicting aggregate uncertainty and aggregate unemployment respectively. The figure also plots the observed D_{dpard}, which is the difference in the prediction criteria test statistic (red line) and $d^*_{dpard}(\alpha)$, which is the left tail of the simulated distribution of the prediction criteria test statistic (dotted blue line) and assesses the ability of the disaggregate model versus the partial disaggregate model without state-level uncertainty in predicting unemployment.

Figure A11: Prediction Criteria - Robustness Analysis using GDP Weights



Note: This figure compares the IRFs of an aggregate uncertainty shocks in the GVAR model to the VAR model. The GVAR IRFs are the weighted response of each state unemployment to an aggregate uncertainty shock in the GVAR model, where the aggregate shock is defined in Equation 17 through using population weights. The VAR IRFs are constructed according to Equation 8 via the Cholesky decomposition where uncertainty is placed first. Unemployment is defined as the quarterly change in unemployment rate. The confidence interval is the bootstrapped IRFs at ± 1 s.d.

Figure A12: Aggregate Uncertainty Shocks on Aggregate Unemployment - Robustness Analysis using GDP Weights



Notes: This figure plots $\frac{p_{ii}^{NU}}{p_{ii}^{U}}$, which measures the relative importance of national influences in propagating uncertainty shocks on state i's unemployment (the blue dot). The 95 % bootstrapped confidence interval are denoted by the dotted lines.

Figure A13: Relative Importance of National Uncertainty - Robustness Analysis using GDP Weights