The Value of News

By Vegard H. Larsen and Leif A. Thorsrud*

We decompose a major business newspaper according to the topics it writes about, and show that the topics have predictive power for key economic variables and, especially noteworthy, for asset prices. Unexpected innovations to an aggregated news index, derived as a weighted average of the topics with the highest predictive scores, cause large and persistent economic fluctuations, a permanent increase in productivity, and are especially associated with financial markets, credit and borrowing. Unexpected innovations to asset prices, orthogonal to news shocks and labeled as noise, have only temporary positive effects. (JEL C8, D84, E32, O33)

There is a widespread belief that changes in expectations, due to news, could be an important independent driver of economic fluctuations. In modern research, a commonly applied mechanism for rationalizing this belief formulates the expectation formation process as a signal extraction problem, see, e.g., Beaudry and Portier (2014): At each point in time the agents in the economy receive a signal about the economy's future needs and developments. One part of the signal is true news representing fundamental information; the other is noise. When the agents manage to filter a positive signal correctly and act accordingly, the economy booms. When the agents respond positively to a signal that turns out to be noise, the economy initially booms, but then contracts as the agents

*Larsen: Centre for Applied Macro and Petroleum economics, BI Norwegian Business School, and Norges Bank (email: `vegard.h.larsen@bi.no`); Thorsrud: Centre for Applied Macro and Petroleum economics, BI Norwegian Business School (email: `leif.a.thorsrud@bi.no`). This paper is part of the research activities at the Centre for Applied Macro and Petroleum economics (CAMP) at the BI Norwegian Business School. We thank George-Marios Angeletos, Hilde C. Bjørnland, Fabio Canova, Gernot Doppelhofer, Ragnar Juelsrud, Gisle Natvik, Kjetil Storesletten, Tommy Sveen, and Mike West for valuable comments. The views expressed are those of the authors and do not necessarily reflect those of Norges Bank. The usual disclaimer applies.

revise their expectations. True news is therefore information that should have predictive power for the future developments in the economy. However, for the researcher trying to measure the macroeconomic effects of news there is a challenge: New information is not observed.

Following the pioneering work of Beaudry and Portier (2006), empirical investigations have tried to solve the unobservable problem by building on a basic tenet in finance, namely that asset prices change in response to unexpected fundamental information. However, using changes in asset prices to measure news does not permit the researcher to identify the type of fundamental information that actually causes the prices to change. It is also well documented that asset prices tend to under- or over-react to new information, depending on the circumstances, see, e.g., Tetlock et al. (2008). Thus, using unexpected innovations in the stock market as a proxy for news shocks is not a satisfying solution.[1] Moreover, the financial literature itself has not been able to document a robust relationship between stock prices and news, in its literal meaning, although, as argued by Boudoukh et al. (2013), this might merely be because the literature has used the wrong measures of news.

In this paper we construct a novel and more direct measure of news - namely one that is based on what is actually written in the newspaper. More precisely, we decompose a major business newspaper according to the topics it writes about using a Latent Dirichlet Allocation (LDA) model. The LDA model statistically categorizes the corpus, i.e., the whole collection of words and articles, into topics that best reflect the corpus's word dependencies.[2] A vast information set consisting of words and articles can thereby be summarized in a much smaller set of topics facilitating interpretation and usage in a time series context.

Our main motivation is to address the predictions given by the news driven business cycle view. To do so we continue by investigating which news topics

[1] To circumvent some of the issues related to using the stock market to identify news innovations, some studies instead use unexpected innovations in consumer confidence, with Barsky and Sims (2012) being a primary example. Still, and as for asset prices, such innovations do not say anything about the type of fundamental information that actually constitutes news.

[2] Blei et al. (2003) introduced this model as a natural language processing tool. Since then the methodology has been heavily applied in the machine learning literature and for textual analysis. Surprisingly, in economics, it has hardly been applied. See Hansen et al. (2014) for an interesting exception.

predict (macro)economic outcomes and derive an aggregated news index based on these results. We then use the news index together with asset prices in Structural Vector Autoregressive (SVAR) models to identify news and noise shocks. Following Beaudry and Portier (2006), news shocks are restricted to be orthogonal to unanticipated productivity shocks, while we treat the variation in asset prices not explained by unanticipated productivity and news shocks as noise innovations. Our hypothesis is simple: To the extent that the newspaper provides a relevant description of the economy, the more intensive a given topic is represented in the newspaper at a given point in time, the more likely it is that this topic represents something of importance for the economy's future needs and developments. Thus, instead of relying on innovations in the stock market to measure news, we use a primary source for news directly - the newspaper.

Our analysis adds to the literature along two related fronts. First, our analysis adds to the literature investigating the empirical importance of news and noise shocks. We refer to Beaudry and Portier (2014) for an excellent overview of the current strand of both the theoretical and empirical aspects of this literature.[3] We contribute to this research agenda in entertaining a more direct measure of news, namely news topics. We argue that this allows us to empirically investigate the macroeconomic effects of news and noise shocks which are key in the theoretical mechanism used to rationalize the news driven view of the business cycle.[4] Moreover, unlike existing methodology, our approach allows us to identify the type of new information (in terms of news topics) that actually constitute a news shock.

Second, our approach shares many features with a growing number of studies using textual data to predict and explain economic outcomes, see, e.g., Tetlock (2007), Soo (2013), and Bloom (2014). However, we do not need to subjectively classify the text using negative, positive or specific word counts, as is often done in existing studies. Instead, the LDA machine learning algorithm automatically

[3]A closely related literature studies the role of anticipated shocks as a source of economic fluctuations, see, e.g., Schmitt-Grohe and Uribe (2012). Like news shocks, anticipated shocks are known in advance and contain signals about future economic developments.

[4]The researcher's ability to separately identify these innovations is debated in the literature due to the nonfundamentalness problem, see, e.g., Forni et al. (2014). In essence, our argument rests on the fact that the first stage predictive regressions work as a filtering mechanism for true news. We return to this discussion more fully in Section 3.

delivers topics that describe the whole corpus. Therefore, in contrast to using positive and negative words, the topic based approach permit us to identify the type of new information that might drive economic fluctuations. As argued by Beaudry and Portier (2014), the content of the news could also be about many diverse objects. By employing the LDA decomposition of the news corpus we are, loosely speaking, letting the data speak rather than restricting ourselves to specific word counts. Lastly, what is positive and what negative obviously relates to an outcome. A topic does not. A topic has content in its own right.[5] In relation to this, the news concept we have in mind is linked to fundamental information. By focusing on topics which have a concrete meaning in their own right and potentially predictive power for future economic developments we reduce the "risk" of picking up news that is not linked to fundamentals. For example, using news measures based on positive and negative word counts, news is often interpreted more broadly in line with the classical work of Pigou (1927) and Keynes (1936) on capturing the market's animal spirits where changes in agents' expectation can be totally self-fulfilling or not rooted in economic fundamentals at all.[6]

The empirical application used in this study employs Norwegian text data, collected from Retriever's "Atekst" database, but our methodology for extracting news from newspaper data and classify its predictive power is general. We focus on Norway because small economies, like Norway, typically have only a few business newspapers, making the choice of corpus less complicated. Here, we simply choose the corpus associated with the largest and most read business newspaper, *Dagens Næringsliv* (DN), noting that DN is also the fourth largest newspaper in Norway irrespective of subject matter.[7]

[5]In relation to an outcome, however, the sign of the topics matter. As described in Section 2, we statistically identify whether the news topics signal positive or negative news at a given point in time using a time-varying parameter model when performing the predictive regressions.

[6]This does not mean we think such mechanisms are unimportant. In fact, a growing literature suggests they are, see, e.g., Beaudry et al. (2011) and Angeletos and La'O (2013). Moreover, ex-ante, we cannot guarantee the news topics extracted using the LDA model do not capture such mechanisms. It is plausible that items in the newspaper generate a self-fulfilling feedback loop where the mood of the news changes economic activity, thus validating the original sentiment. This caveat, however, is not limited to our study, it applies to all empirical studies in this literature.

[7]In addition, Norway is a small and open economy and thereby representative of many western

Our main results are as follow. First, the decomposition of the DN corpus into news topics using the LDA model produces topics that are easily classified; it describes the DN corpus well statistically, but also intuitively for those with knowledge of DN and the Norwegian economy. A wider range of these news topics adds marginal predictive power. This holds particularly for output, but also for forward-looking variables such as business sentiment and asset prices. The latter finding is particularly noteworthy as the financial literature has produced little evidence of a link between news and returns; see the discussion in Boudoukh et al. (2013).

Second, irrespective of whether we estimate bivariate SVARs, as in Beaudry and Portier (2006), or larger systems entertaining both the news index and asset prices, unexpected news innovations cause large and persistent economic fluctuations and a permanent increase in productivity, in line with existing empirical evidence. In contrast to existing studies, however, we show that the news shocks are particularly related to news topics describing developments in the financial markets, credit and borrowing; but many other topics make significant contributions. Among these, and especially important in the Norwegian economy, are topics associated with the energy sector.[8]

Third, when specifying a SVAR including both the news index and asset prices, we are able to confirm the main predictions from prominent theoretical news driven business cycle models, see, e.g, Barsky and Sims (2012) and Blanchard et al. (2013): Unexpected innovations in the news index cause (i) a fall in inflation and a rise in the real interest rate; and (ii) a persistent increase in consumption, employment, hours and TFP. On the other hand, after unexpected innovations to asset prices, orthogonal to news shocks and labeled as noise, (iii) consumption, employment, and inflation rise for a short time period, only to fall back again. Thus, news and noise shocks operate very much like supply and demand shocks, respectively. Together, (iv) the two shocks explain a non-negligible share of the long-run economic fluctuations in consumption and

countries. DN was founded in 1889, and has a right-wing and neoliberal political stance.

[8]As such, our findings encompass some of the results presented in, e.g., Ramey and Shapiro (1998), Romer and Romer (2010), Mertens and Ravn (2012), Dominguez and Shapiro (2013), and Arezki et al. (2015), which provide concrete, independent, examples where anticipated shocks, or news, are linked to expectations about future policy, energy prices, and industrial explorations. However, the methodology employed in these papers differs markedly from ours.

productivity, and almost 100 percent of the short-run variation in asset prices.

An implication of our findings is that models that identify news shocks using asset prices are likely to confound the effects of news and noise shocks. For this reason, our interpretation of a news shock does not accommodate the ones typically described in the empirical literature.[9] However, as mentioned above, the results listed in (i) - (iv) do resemble those obtained in prominent theoretical news driven business cycle models. In these models news is restricted to work through a productivity channel directly (as anticipated productivity shocks). Our finding that a broad range of news topics actually contribute significantly to news shocks calls into question the validity of such a restriction, but suggests that it's not a bad approximation. Alternatively, our findings should be suggestive for future work on how news shocks theoretically transmit and ultimately affect productivity and economic fluctuations.

The rest of this paper is organized as follows. In Section 1 we describe the newspaper data, the topic model, and the estimated news topics. We describe how we construct an aggregated news index in Section 2. In Section 3 we present the SVAR experiment. Section 4 includes additional results and a discussion of implications. Section 5 concludes.

1 The News Corpus and the LDA

The *Dagens Næringsliv* (DN) news corpus is extracted from Retriever's "Atekst" database, which gives access to all articles published in DN from May 2 1988 to December 29 2014. We retrieved a total of 459 745 articles, well above one billion words and more than a million unique tokens, covering a sample of over 9000 days. This massive amount of data makes statistical inference challenging, but as is customary in this branch of the literature we take some steps to clean and reduce the raw dataset before estimation.

We start by filtering out words from a stop-word list. This is a list of common words we do not expect to have any information relating to the subject of an article. Examples of such words are *the*, *is*, *are*, and *this*. We also remove the most common Norwegian surnames and given names. In total the stop-

[9]See, in particular, Beaudry and Portier (2006) and Barsky and Sims (2011) for two contrasting interpretations, and the discussion in Section 4.

word list together with the list of common surnames and given names removed roughly 1800 unique tokens from the corpus. Next, we run an algorithm known as stemming. The objective of this algorithm is to reduce all words to their respective word stems. By word stem we mean the part of a word that is common to all of its inflections. An example is the word *effective* whose stem is *effect*. The last thing we do is to calculate a corpus measure called *tf-idf*, which stands for term frequency - inverse document frequency. This measures how important all the words in the complete corpus are in explaining single articles. The more often a word occurs in an article, the higher the *tf-idf* score of that word. On the other hand, if the word is common to all articles, meaning the word has a high frequency in the whole corpus, the lower that word's *tf-idf* score will be. We keep around 250 000 of the stems with the highest *tf-idf* score, and use this as our final corpus.[10]

To quantify the value of news we start by decomposing the DN corpus according to the topics it writes about using a topic model. In general, topic modeling algorithms are statistical methods that analyze the words of the original texts to discover the themes that run through them and the themes connection to each other. One of the simplest topic models, and the one used here, is the Latent Dirichlet Allocation (LDA) model. The LDA model is an unsupervised learning algorithm introduced by Blei et al. (2003) that clusters words into topics, which are distributions over words, while at the same time classifying articles as mixtures of topics.[11] By unsupervised learning algorithm we mean an algorithm that can learn/discover an underlying structure in the data without the algorithm being given any labeled samples to learn from. The term "latent" is used because the words, which are the observed data, are intended to communicate a latent structure, namely the meaning of the article. The term "Dirichlet" is used because the topic mixture is drawn from a conjugate Dirichlet prior.

[10] We have experimented with different cut-offs for the *tf-idf* score. Moving the cut-off around (within the limits of making the estimation computationally feasible) does not seem to change the results. For stemming we use a Norwegian algorithm implemented in the Natural Language Toolkit (`www.nltk.org`). The stop-word list can be obtained on request.

[11] Blei and Lafferty (2006) and Mcauliffe and Blei (2008) extend the LDA to a dynamic and supervised setting, respectively. Both of these (and other) more advanced extensions are relevant for the problem addressed in this paper. We leave it to future research to assess their merit in doing so.

At an intuitive level, the best way to understand the LDA model is to start by making a thought experiment of how the articles in the newspaper were generated. Following Blei (2012), assume we know all the topics, then the procedure by which articles are generated will be as follows:

1. Pick the overall theme of an article by randomly giving it a distribution over topics
2. For each word in the document
   (a) From the topic distribution chosen in 1., randomly pick one topic
   (b) Given that topic, randomly choose a word from this topic

Iterating on 2) generates an article that will possibly consist of many different topics, but where one of these is more important than the others. Iterating on 1) and 2) generates a large set of articles, each of which might be distinguished by which topics best describe this article.

The technical details on estimation and prior specification for the LDA model are described in Appendix G.1. Here we note that we estimate the model using 7500 Gibbs simulations and classify 80 topics. In Table 4, in Appendix F, we show that this latter choice is preferred on statistical grounds, meaning that 80 topics provide a good statistical decomposition of the DN corpus. We have experimented with using fewer topics, observing that our main results do not change. Chang et al. (2009) document that topic models which perform better on statistical criteria may infer less semantically meaningful topics (relative to how humans interpret the corpus). Below we document that the decomposition of the corpus using 80 topics results in interpretable objects.

1.1 News Topics

Table B, in Appendix B, lists all the estimated topics together with the most important words associated with each topic. The LDA estimation procedure does not give the topics any name or label. Still, as seen from the table, it is, in most cases, conceptually simple to classify the topics. Although there is no correct way of doing this, we believe most people would more or less agree with our approach, see the second and fourth column of each table, and the discussion in Appendix B. That said, the labeling plays no material role in our experiment, it just serves as a convenient way of referring to the different topics

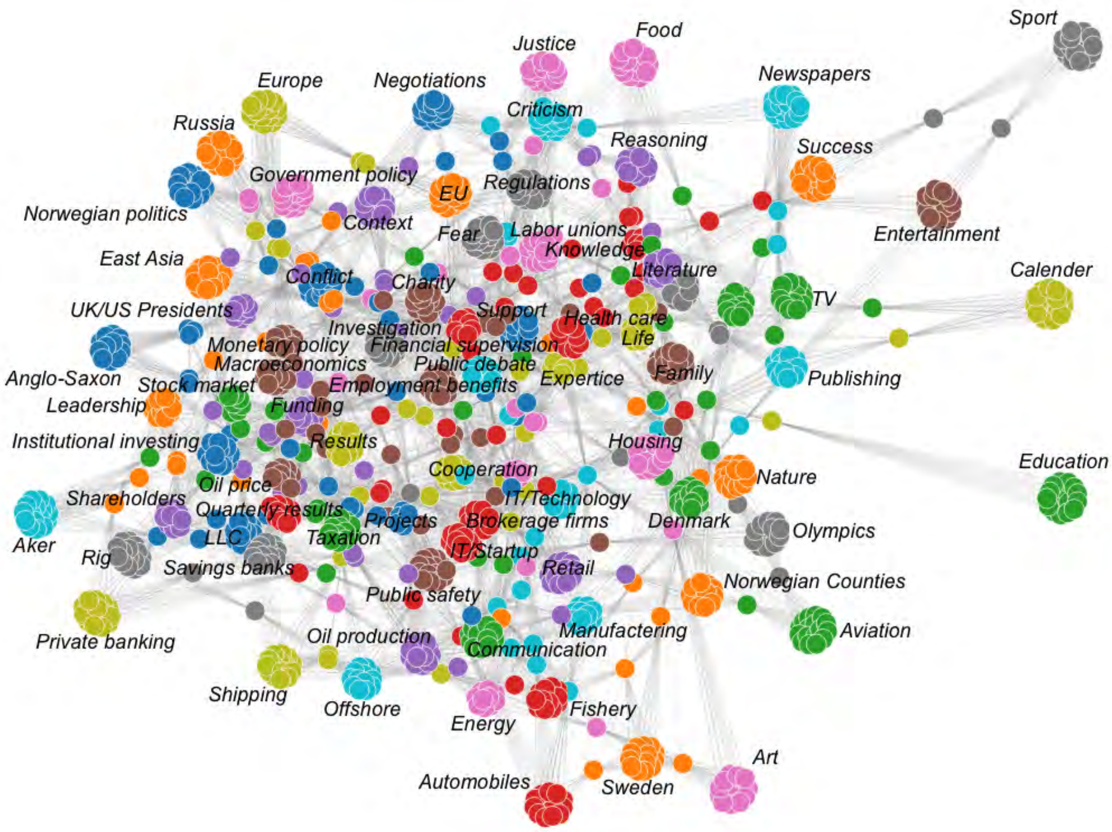


Figure 1. DN visualized using a topic nett. Different colors are used for words in different topics. Words that belong to the same topic have an edge between them.

(instead of using, e.g., topic numbers or long lists of words). What is more interesting, however, is how the words within and between the different topics relate to each other. Figure 1 addresses this issue.

Figure 1 gives a network representation of the estimated topics. The topic net is created as follows. We select the 17 most important words in each topic and give these words the same color.[12] These words are the nodes in the graph. For readability we do not label the nodes, only the topic's name using the subjective classifications discussed above and in Appendix B. The edges connecting words across topics show the degree to which the same words occur in different topics. For example, the *Sport* topic shares one word with the *Entertainment* topic, and one word with the *Success* topic. On the other hand, a topic like

[12]Seventeen words were chosen for visual clarity. A larger number of words increases the complexity of the graph, making it harder to see the different topics. When a word belongs to more than one topic the color assigned to that word is arbitrarily selected to match the topic with the lowest number.

Taxation shares words with many other topics.

An important message from the decomposition reported in Figure 1 is that the same words often occur in different topics. Actually, when using the first 17 words of each topic, all the 80 topics share at least one word with another topic. Thus, topics cluster together because they share words, indicating that they also relate in theme and meaning. For example, as we see from Figure 1, topics such as *Energy* and *Oil Production* stand close to each other. So do topics associated with *Savings Banks*, *Shareholders* and *Institutional investing*. On the other hand, some topics, like *Education*, *Sport*, *Art* and *Newspapers*, seem more isolated. This clustering can be easily explained if we know how DN structures its content, with distinct sections for, e.g., media and art. Finally, although many of the topics reported in Figure 1 are relatively general, many of them make it clear that DN is a Norwegian newspaper writing about news of particular relevance for Norway. We observe separate topics for Norway's immediate Nordic neighbors (*Sweden* and *Denmark*); largest trading partners (*EU* and *Europe*); and biggest and second biggest exports (*Oil production* and *Fishing*).

Given knowledge of the topics (and their distributions), we translate the decomposition of the DN corpus into time series that can be used to evaluate the value of news in explaining economic fluctuations: For each day we calculate the frequency with which each topic is represented in the newspaper that day. By construction, across all topics, this number will sum to one for any given day. On average across the whole sample, each topic will have a more or less equal probability of being represented in the newspaper. However, across shorter time periods, i.e., months or quarters, the variation can be substantial. This is illustrated in Figure 2, which reports the time series for two of the 80 topics estimated.[13] As is clearly visible in the figure, the time series for each topic fluctuates substantially across time.

In each graph in Figure 2, we also report a measure of the Norwegian business cycle. By simple visual inspection we observe that the *Funding* topic covaries with this measure, at least during specific time periods. Also, the *Fear* topic seems to capture important business cycle swings, but misses the timing more

[13]The numbers are reported here and throughout this paper on a quarterly frequency. The aggregation from day to quarter is obtained as a simple mean.

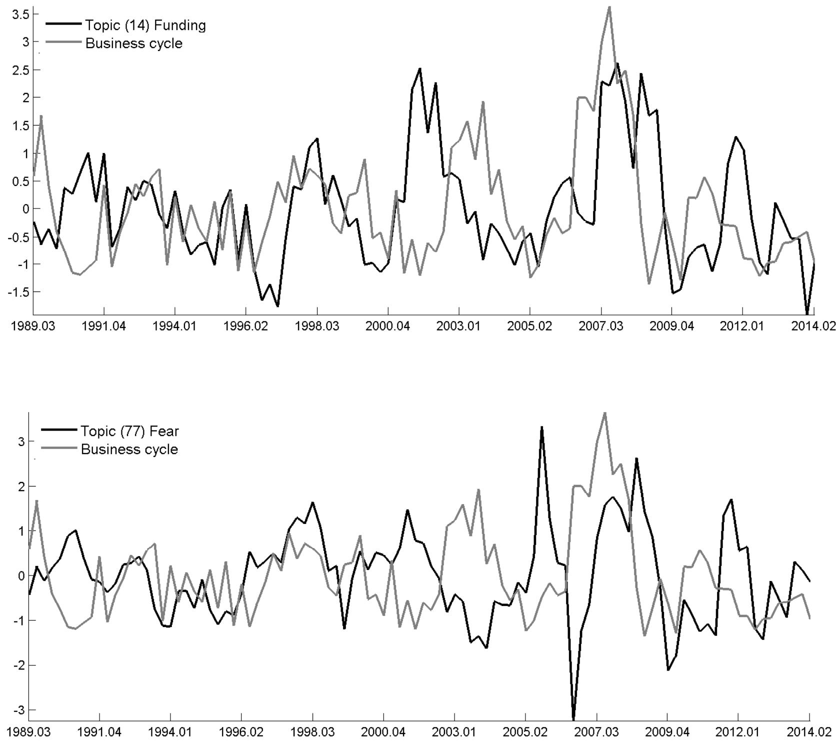


Figure 2. Each individual topic time series is transformed to year-on-year growth rates and standardized. The business cycle measure is the standardized value of Hodrick Prescott (HP, $\lambda = 40000$) filtered GDP. Because the topics are not sign identified, see Section 2, the business cycle estimate is reported in absolute value.

often than the former. It is also evident from the figure that some of the topics might be correlated with each other. The maximum (minimum) correlation across all topics is 0.66 (-0.44). However, overall, the average absolute value of the correlation among the topics is just 0.1, suggesting that different topics are given different weight in the DN corpus across time.

2 Constructing an aggregate news index

As alluded to in the introduction, for news to have an important role in explaining economic fluctuations, it needs to predict economic outcomes.[14] We analyze the predictive power of the news topics by running a battery of predic-

[14]Naturally, the news topics described in the previous section can simply be news about yesterday's events, and not forward looking at all. That said, if the economic agents receive news on events in the past, their expectations about the future may still be affected.

tive regressions for a number of outcome variables, central in the news driven business cycle literature: output (Y); investment (I); consumption (C); total factor productivity (TFP); asset prices ($OSEBX$); and business confidence (BCI).[15]

The goal of the predictive experiment is not to run a horse race with different predictors and model specifications to obtain the specification that delivers the best fit. Instead, by comparing the predictive power of the different news topics, the predictive regressions essentially filter out topics which on average contain forward looking information, i.e., have marginal predictive power, from those news topics that do not. It thereby facilitates the construction of an aggregated news index useful for business cycle analysis. In addition, we use the predictive regressions to identify the sign of the news topics. We describe the specification, estimation, and scoring algorithm we employ to do so in greater detail below. The aggregated news index is presented in Section 2.3.

2.1 Specification, estimation, and scoring

For each outcome variable, the predictive regressions are specified as AR(p) or ARX(p) regressions, i.e., autoregressive processes of order p, with a topic (ARX) or without a topic (AR) included. Accordingly, for a given outcome variable, we run T number of ARX(p) models, where each ARX(p) is differentiated by the topic it includes.

We estimate both the AR(p) or ARX(p) specifications using a Latent Threshold Model (LTM). The LTM was introduced by Nakajima and West (2013), and can be written as follows:

$$y_t = x'_{t-1} b_t + u_t \qquad u_t \sim N(0, \sigma_u^2) \tag{1a}$$

$$b_t = \beta_t \varsigma_t \qquad \varsigma_t = I(|\beta_t| \geq d) \tag{1b}$$

$$\beta_t = \beta_{t-1} + e_t \qquad e_t \sim N(0, \Sigma_e) \tag{1c}$$

where t is the time index, x_{t-1} is a $(n \times 1)$ vector of variables used for prediction, b_t a $(n \times 1)$ vector of time-varying parameters. ς_t is a zero one variable, whose value depends on the indicator function $I(|\beta_t| \geq d)$. If the ith element in $|\beta_t|$ is

[15] The BCI measure is used because no consumer confidence measure exists on a quarterly frequency covering the sample entertained here. Additional details about the data are described in Appendix A.

above the ith element in the $(n \times 1)$ threshold vector d, then $\varsigma_t = 1$, otherwise $\varsigma_t = 0$. e_t is a $(n \times 1)$ vector of disturbances associated with the time-varying parameters. We assume that e_t and u_t are independent.

Note that in equation (1a), the whole x_{t-1} vector is lagged one period relative to y_t. For the analysis conducted in Section 3 this is important. In particular, for the ARX(p) models we avoid simultaneity issues resulting from the fact that the dependent variable in the regression might (within the quarter) affect what is written about in the business newspaper, and hence, the newspaper topics.

In general, the LTM is a useful estimation strategy for models where the researcher wants to introduce dynamic sparsity into the system. In our context, the LTM serves two purposes. First, the time series for each topic will be an intensity measure. While the sign of this measure in relation to an outcome variable is not identified from the topic extraction itself, the time-varying parameter formulation used in (1) allows us to identify the sign of the news in relation to an outcome variable. If estimating a predictive regression like (1a) without time-varying parameters, the researcher might conclude that a given topic has no predictive power for y_t, i.e., that $b = 0$, simply because, on average, periods with a positive b_t cancels with periods with a negative b_t. By using the time-varying parameter formulation in (1), we avoid this pitfall. Second, by introducing the threshold dynamics, we also safeguard against over-fitting. When running T predictive regressions for each outcome variable, some topics might, by chance, be classified as having marginal predictive power. Enforcing a threshold reduces this risk. Moreover, it is not very likely that one particular topic is equally important throughout the estimation sample. A topic might be very informative in some periods, but not in others. The threshold mechanism potentially captures such cases in a consistent and transparent way.

The system in (1) is estimated using Gibbs simulations. The details, together with prior specifications, are described in Appendix G.2. We set $p = 1$, and the estimation sample is 1988:Q3 – 2014:Q4. Y, I, C, TFP, $OSEBX$, and the topics are all transformed to year-on-year logarithmic differences, $y_t = ln(Y_t) - ln(Y_{t-4})$, before estimation.[16] The BCI indicator is used in levels. To reduce

[16]The transformation is done to ensure the topics are stationary. In theory it would be hard to imagine that the news topics would be anything but stationary: the constructed news topic time series are bounded between 0 and 1 by construction. Still, across a limited estimation sample, non-stationary topics might be observed. We have also experimented with using the

the impact of potentially tilting the priors toward a given explanatory variable, all variables are standardized.

In a Bayesian setting, the natural scoring metric is the marginal likelihood of model M_i relative to M_j for $i \neq j$, i.e., the posterior odds ratio. The marginal likelihood for model i can be written as:

$$p(y|M_i) = \int p(y|\theta^i, M_i)p(\theta^i|M_i)d\theta^i, \tag{2}$$

where θ^i are the parameters of the model, $p(y|\theta^i, M_i)$ is the likelihood, and $p(\theta^i|M_i)$ is the prior. Under equal priors for each model, which we assume, the posterior odds ratio is given by:

$$PO_{ij} = \frac{p(y|M_i)}{p(y|M_j)}. \tag{3}$$

We note that the posterior odds ratio will favor models with a better fit, in the sense that if favors models with less residual variance, but also models where there is greater coherency between the prior and the information in the data.

In presenting the results below, we treat the j model specification to compare against as the AR(p) specification, while the ARX(p) specification, for $i = 1, \ldots, T$, are the alternatives. A higher value of PO_{ij} implies a higher posterior probability for model i relative to model j, i.e., evidence that the topics add marginal predictive power for y_t beyond whatever is captured in y_{t-1} itself.

2.2 The value of news in prediction

Figure 3 summarizes the predictive results. The plot reports all topics, and associated outcome variables, where the posterior odds ratio is $lnPO_{ij} > 2$. In a Bayesian setting, see Kass and Raftery (1995), such model scores are assumed to represent good evidence in favor of model i relative to model j. In the figure, a thicker line connecting a given topic and outcome variable signals a higher value of $lnPO_{ij}$ (over and above 2).

The first finding is that irrespective of which variable is being predicted many topics actually add marginal predictive power. Still, the most predictable variable by far, in terms of using news topics, is output. Almost all topics listed

news topics in levels and using trend (linear) adjusted news topics in the predictive regressions. Qualitatively, our main results are not affected by these choices.

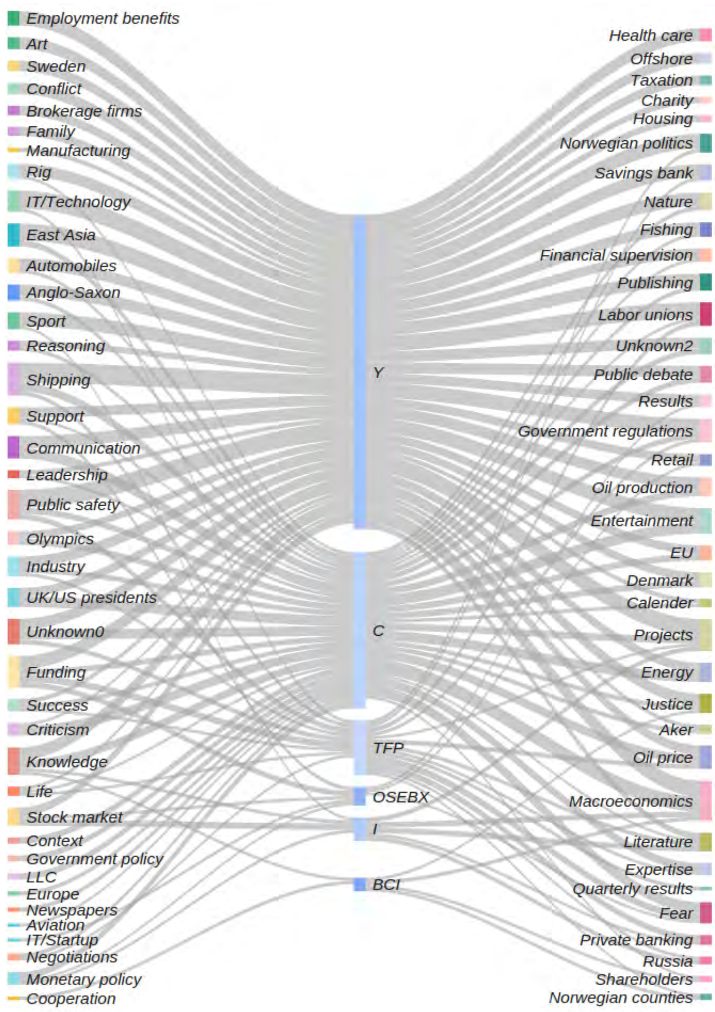


Figure 3. The Sankey diagram reports all topics, and associated outcome variables, where the posterior odds ratio is $lnPO_{ij} > 2$. A thicker line indicates that this relative score is higher (over and above 2).

in the figure have a connection with output, and many of the lines connecting the topics to output are relatively thick, suggesting a high posterior odds ratio and strong evidence in favor of the topic augmented regressions. Similar results, although not as strong as for output, are found for consumption.

The second striking finding is that for forward looking variables like asset prices and business confidence, supposed to contain all fundamental information already known in the economy, the topic augmented regressions also seem to add marginal predictive power. Given the lack of evidence in the financial literature that news predicts asset prices, this is surprising. Interestingly, similar results have recently been documented by Boudoukh et al. (2013). According to them, the lack of a predictive linkage between news and asset prices might simply be because the literature has been employing bad measures of news. For this reason, Boudoukh et al. (2013) also classify news into topics and find that news actually helps predict returns. However, although we reach similar conclusions, the methodology and experiment conducted by Boudoukh et al. (2013) is very different from ours.[17]

So, do the news topics that add marginal predictive power also make sense from an economic point of view? We believe they do. For example, we see from Figure 3 that the *Shipping* topic gets a relatively high score in predicting output. So do topics such as *Oil production* and *East Asia*. As Norway is a small and open petroleum exporting economy, with a banking sector highly oriented toward shipping, this is reasonable. Moreover, the *Macroeconomics* topic receives a high score in predicting consumption and business confidence, the *Stock Market* topic is important for investments, and the *Funding* topic is important for both asset prices and productivity. Still, some news topics that receive a high score might, on face value, seem to reflect a spurious relationship. A case in point is the *Literature* topic, which adds marginal predictive power to the regressions for investments. That said, such exotic relationships are the exception rather than the rule. It is also the case that a given newspaper article contains many topics at the same time. To the extent that different topics, meaningful or not from an economic point of view, stand close to each other in the decomposition of the corpus, see Figure 1, they might covary and both might therefore add value in terms of predicting economic outcomes.

Finally, one might be surprised by the widely varying scores of the different topics in explaining the different outcome variables. One important reason for this is because the predictive regressions differ across outcome variables. For example, when predicting asset prices using the ARX(p) model, we condition on lagged asset prices in addition to a topic. When predicting output using the same model, we condition on lagged output in addition to the topic. If the informational content of the lagged dependent variable differs markedly across these predictive regressions, so will the weight given to the different topics. Thus, although the *Macroeconomics* and *Stock Market* topics seem

[17] In relation to this, when we compare the predictive performance of our topic based approach to a keyword based approach, based on counting positive and negative words, we find that the topic based approach seems superior. This experiment and associated results are described in greater detail in Appendix C.

of less importance for asset prices, which is surprising, the information they contain is very likely already captured in lagged asset prices themselves, and is therefore not news in itself.

2.3 The news index

To construct an aggregate news index we use the predictive results already obtained. We focus on news that predicts asset prices, i.e., $OSEBX$. The reason is threefold: First, economic theory dictates that asset prices should contain all available fundamental information relevant to the economy. News topics that receive a large weight in predicting asset prices should therefore also reflect the most important fundamental information relevant to the economy. Second, unlike many other economic variables like, e.g., output and consumption, asset prices are not revised. If they had been subject to revision we would have needed to make a choice regarding which vintage of data to use to best reflect the state of the economy at each point in time. Third, using asset prices has been a guiding principle in the existing (empirical) news driven business cycle literature.

We do, however, acknowledge that the use of other outcome variables than asset prices when constructing the aggregate index can be defended. As alluded to in the introduction, and as documented in a large body of literature, motivated through work in behavioral finance and limits of arbitrage, stock prices tend to under- or over-react to news, depending on the circumstances, see, e.g., Vega (2006), Gutierrez and Kelley (2008), and Tetlock et al. (2008).[18] Therefore, if the aggregated news index ends up as a perfect fit of the observed asset prices, we might have gained very little in terms of constructing an index reflecting true news about fundamentals. As it turns out, and discussed more fully below, this is not the case. The constructed news index does not fit asset prices perfectly. Another possibility would be to use, e.g., TFP, to identify more clearly an aggregated news index with predictive power for future productivity developments. Although common in the literature, we do not want to restrict ourselves to such an interpretation. As exemplified by numerous studies, see, e.g., Romer and Romer (2010), Mertens and Ravn (2012), and Arezki

[18]This also provides one likely explanation for the varied usefulness of using financial market information in predicting macroeconomic outcomes, see, e.g., Stock and Watson (2003).

et al. (2015), the news that agents act upon can potentially be news about many diverse objects such as; future policy, energy prices, and technological developments, since any of these objects will affect the economy's future needs and development. On the other hand, there are no restrictions in our topic estimation or aggregation that restrict the index to be anything other than news about future productivity either. In addition, productivity is not an observable variable, and has to be estimated. We are skeptical to constructing an aggregated news index based on both an estimated outcome variable and estimated news topics.[19]

Formally, we construct the aggregated news index, denoted NI_t, for each time period t, based on the following weighting formula:

$$NI_t = \sum_{i=1}^{T} w_i b_{i,t} n_{i,t-1}, \tag{4}$$

where $n_{i,t-1}$ is topic i at time $t-1$, and $b_{i,t}$ is the estimated parameter (from the LTM model) for topic i at time t. w_i is the weight attached to topic i in predicting y_t in equation (1a), constructed using the marginal likelihoods from each predictive ARX(p) model such that:

$$w_i = \frac{p(y|M_i)}{\sum_{i=1}^{T} p(y|M_i)}. \tag{5}$$

Thus, the weights sum to one. For brevity they are reported in Figure 12, in Appendix F. We note that although many news topics add marginal predictive power, confer Figure 3, only a few topics receive a large weight. We discuss these more fully in Section 3.4.

Our preferred measure of the aggregated news index is reported in Figure 4. As is clearly seen in the plot, although we have weighted the topics according to how well they predict asset prices, the aggregated news index does not resemble asset prices perfectly. The aggregated news index often lead the major movements in asset prices, moreover. This follows naturally from how we have constructed the news index, namely as a weighted average of those topics that predict future asset prices. We also see from the plot that the news index seems

[19] Despite this, and as discussed in Section 4, our main results are robust to using an aggregated news index based on how well the topics predict TFP. One reason for this is likely because the news topics that receive a large weight when predicting asset prices and productivity are very much the same, see Figure 12, in Appendix F.

Figure 4. The aggregated news index, *OSEBX* and a measure of the business cycle. The business cycle measure is the Hodrick Prescott (HP, $\lambda = 40000$) filtered GDP. All series are standardized.

to lead the general business cycle, at least at certain points in time. For example, prior to the Great Recession, the news index turned negative well before the turn of the cycle. The news index also signaled the start of the boom during the mid 2000s, well ahead of time.[20]

3 The value of news in business cycles

The main motivation for our analysis is to use textual data to address the predictions given by the news driven view of the business cycle. To do so we use a Vector Autoregressive (VAR) model.[21] We differ from previous empirical studies in that we exchange asset prices with our proposed news index, but still include both variables in the VAR. This allows us to identify both news and noise shocks and, equally important, investigate what type of news actually constitute a news shock.

Below we elaborate on these points. We first provide a more formal, yet

[20] The leading properties of the aggregated business cycle index are further documented in Figure 13 in Appendix F. The figure reports the empirical autocorrelation, at lead and lags, between the aggregated news index and the four key macroeconomic variables output; investment; consumption; and productivity. For output and investments the news index seems to be particularly leading, with a significant autocorrelation coefficient even for the 6th lag, while none of the results on either of the macro variables seems to suggest that news is lagging.

[21] The VAR is a workhorse model for doing empirical macroeconomics. By employing a suitable mapping between reduced form residuals and structural shocks, causal inference can be conducted making the VAR structural, i.e., a SVAR. See Appendix G.3 for technical details, and Beaudry and Portier (2014) for an overview of its usage in the news literature.

simple, description of how the signal extraction problem faced by the agents in the economy can be modeled in a theoretical setting, confer the very first paragraph of this article, and how it has been difficult to validate empirically due to the problem of nonfundamentalness. We then show how we overcome this problem using our news index in combination with asset prices. The empirical evidence brought forward by employing our proposed identification method is presented in Section 3.3, while the decomposition of the news shock into topic contributions is discussed in Section 3.4.[22]

3.1 Some theory and the nonfundamentalness problem

As is customary in this branch of the literature, assume that the dynamic process for productivity is exogenous and that the agents in the economy only observe a noisy signal of true news. In particular, following the exposition in Forni et al. (2014) closely, let the productivity process be:

$$a_t = a_{t-1} + \epsilon_{t-1} \tag{6}$$

where ϵ_{t-1} is the news shock dated with a lag to reflect that it is anticipated, and:

$$s_t = \epsilon_t + \eta_t \quad \eta_t \sim i.i.d.N(0, \sigma_\eta^2) \text{ and } \epsilon_t \sim i.i.d.N(0, \sigma_\epsilon^2) \tag{7}$$

describe the noisy signal observed by the agents at time t, with the news (ϵ_t) and noise (η_t) disturbances being uncorrelated.[23]

The key ingredient in the signal extraction mechanism described here is that expectations are formed on the basis of a limited information set. The delayed effect of the news shock in affecting a_t means that the information set available to the agents at time t is not sufficient to distinguish the current true news shock from the noise component. However, at time $t+1$ the agents learn about yesterday's news because $\Delta a_{t+1} = \epsilon_t$.

The consequences of this can be illustrated by looking at how real variables respond to news and noise innovations. To do so we continue with some simplistic, but illustrative, assumptions: Agents set consumption, c_t, on the bases

[22] Appendix D provides a comparison between our proposed method for identifying news shocks and more traditional methods.

[23] Since our focus is on news and noise shocks, the process in (6) is deliberately kept simple. See, e.g., Barsky and Sims (2012) and Blanchard et al. (2013) for more sophisticated processes.

of expected long-run fundamentals, output, y_t, is fully demand determined, and employment adjusts to clear the labor market. Thus, $c_t = y_t$, and:

$$c_t = E(a_{t+j}|I_t) = E(a_{t+1}|I_t) = a_t + E(\epsilon_t|I_t) \tag{8}$$

where the equalities follow from the assumed process for a_t given in (6), see Forni et al. (2014) for details, and $E(\epsilon_t|I_t)$ reflects the agents signal extraction problem. In a linear and Gaussian setting, like the one described here, it can be optimally solved using the updating equations associated with the Kalman filter. Since lags of a_t and s_t are uninformative about ϵ_t, $E(\epsilon_t|I_t)$ is the projection of ϵ_t on s_t:

$$E(\epsilon_t|I_t) = \gamma s_t \tag{9}$$

where $\gamma = \sigma^2_\epsilon/\sigma^2_s$ is the signal to noise ratio with $\sigma^2_s = \sigma^2_\eta + \sigma^2_\epsilon$. Combining (8) and (9) we get $c_t = a_t + \gamma(\epsilon_t + \eta_t)$, and the change in consumption is:

$$\begin{aligned} \Delta c_t &= \Delta a_t + \gamma\Delta(\epsilon_t + \eta_t) \\ &= \gamma\epsilon_t + (1-\gamma)\epsilon_{t-1} + \gamma\eta_t - \gamma\eta_{t-1}. \end{aligned} \tag{10}$$

The implications of 10 is that a news shock causes consumption to increase immediately by $\gamma\epsilon_t$, while from the next period and onwards the effect is $y_{t-1}+\epsilon_t$. Thus, news shocks lead to a permanent increase in both consumption, output, and productivity. Conversely, after a noise shock, consumption and output initially booms by $\gamma\eta_t$, but returns to its previous level the following period, while productivity remains unaffected.

The theory model described above highlights the key distinction between how a news shock is assumed to affect the economy relative to a noise shock. The model also exemplifies how the researcher's ability to identify news and noise shocks in empirical settings can be questioned due to the problem of nonfundamentalness. The problem is related to equation (7). If rational agents cannot separate between the news and noise disturbances in real time, the arguments goes, then the econometricians with access to the same information set, will not be able to either. In a VAR setting, this makes it impossible to recover structural news innovations, and noise, as linear combinations of reduced form residuals.[24] For this reason, papers analyzing the empirical relevance of the

[24] Another part of the nonfundamentalness problem relates to the size of the observable information set entertained in the VAR relative to what the agents in the economy use when

news driven business cycle view have almost exclusively relied on quantifying the implications of news and noise shocks using theoretical models that put strong restrictions on the data, as in, e.g., Barsky and Sims (2012) and Blanchard et al. (2013), or by assuming that $\sigma_\eta^2 = 0$ in (7), as in, e.g., Beaudry and Portier (2006). Neither is optimal. The restrictions guided by theory might be questionable and far from reality even though more advanced models than the one described above are used. And, by restricting the signal to be noise free, the problem is just assumed away. As it is well documented that asset prices tend to under- or over-react to new information, depending on the circumstances, this seems like an inadequate solution.[25] We follow a different route.

3.2 Identifying news and noise shocks

In short, our identification approach consists of two steps. In the first step we filter out the component of stock price movements that can be explained by exogenous news topics and construct an aggregate news index. In the second step we use the aggregated news index as an observable variable in the VAR together with asset prices and productivity to separately identify news and noise shocks.

The first step was described in detail in Section 2. In terms of equation (7), we argue that this step can be looked upon as a signal extraction mechanism for true news. In particular, if we treat asset prices as a noisy signal about true fundamentals, the output from the state space system in (1) provides us with news topics that in expectation have the best predictive power for asset prices. At the extreme, if the aggregated news index turned out to explain asset prices perfectly, we would maybe have gained very little in terms of filtering out the relevant information (news versus noise) from asset price movements. However, as shown in Figure 4 already, the news index is not a perfect fit. At the other

making their decisions. In such cases, simply expanding the information set used in the VAR with forward looking variables might solve the problem, see the discussion in, e.g., Watson (1986), Sims and Zha (2006), Forni et al. (2014) and Beaudry and Portier (2014).

[25] Indeed, when we estimate bivariate VARs using various measures of productivity and identify news shocks as innovations to asset prices, as in Beaudry and Portier (2006), we get mixed results. For reasons discussed below, when we exchange asset prices with the news index, we obtain much more robust results. For brevity these bivariate VAR results are presented in Appendix D.

extreme, if the news topics did not predict asset prices at all, this paper would have stranded following Section 2. We can not, however, rule out that the news index we construct contains both anticipated and unanticipated fundamental information.[26] For this, we need step two.

In the second step we use the constructed news index in a VAR, and identify news and noise shocks using a recursive ordering, where we include productivity, the news index, and asset prices, first in the system and in that order. Thus, news shocks are treated as orthogonal to unanticipated contemporaneous innovations in productivity. In line with equation (6), news shocks can therefore be looked upon as anticipated shocks that affect productivity with a delay. Likewise, noise shocks, which we identify as unexpected innovations to asset prices orthogonal to contemporaneous news shocks, can not affect productivity and the news index within the same period. The motivation for ordering asset prices below the news index in the system follows from how we construct the index in step one: It is a linear combination of the news topics that at time $t-1$ best predicts asset prices at time t. Therefore, noise shocks are defined as the component of asset price variation not explained by current fundamental information. Finally, in line with equation (10), by ordering the news index and asset prices above any remaining variables in the system we ensure that these variables are contemporaneously free to move to news and noise innovations.

As in Beaudry and Portier (2006), we put no restrictions on the shock to productivity itself as to allow it to potentially capture a traditional surprise productivity shock, measurement error, or a combination of these. We note, however, that in terms of the theory model described above, such a shock would have made identification difficult. Essentially, the agents would not have been able to separate between past news and surprise productivity shocks when observing the change in productivity. Empirically, when introducing the news index into the VAR, we treat it as a filtered, or "cleaned", observed signal of true news, and thereby avoid the nonfundamentalness problem. The argument

[26] Neither can we rule out that the news index contain only a subset of all fundamental information. As such, our results might be viewed as a lower bound for the importance of news shocks. Moreover, these scenarios are described under the assumption that most of the asset price variation we observe is due to changes in fundamentals. If asset prices are dominated by noise, the news index we construct might end up as actually being a "noise" index. The results presented in Section 3.3 strongly suggest that this is not the case.

rests on the fact that we, as econometricians, have access to more information than the agent's in the economy could have had when making their decisions. To see this, note that ex-ante, or in real time, the news index is not available. The reason is that when constructing the news index in step one an evaluation sample to score the different topics is needed. On the other hand, ex-post, after estimating the predictive regressions, we are able to identify the sign, size, and score of each individual news topic, and use the aggregate index to identify news shocks covering the same sample. A similar argumentation, although using a totally different methodology, is used in recent work by Forni et al. (2014). They show how a signal can be structurally decomposed into news and noise innovations using dynamic identification. Another identification strategy used in the literature, although for news shocks only, was proposed in Barsky and Sims (2011). Compared to our identification strategy, however, their is much more restrictive because it a-priori defines the news shock as the one that maximizes a measure of the forecast error variance of productivity at some horizon.

3.3 Empirical evidence

For all the empirical applications employed here, we specify the VAR with four lags and use the longest estimation sample possible, 1990:Q3–2014:Q2. We consider two different estimates of total factor productivity, labeled TFP and TFP^a. Both measures are based on simple growth accounting and converted into a (log) index, but we correct one of them for variability in capacity utilization (TFP^a).[27] The news index (NI) is used as reported in Figure 4, while asset prices ($OSEBX$) is measured as yearly changes, i.e., $log(x_t) - log(x_{t-4})$. To capture the main business cycle features analyzed in the news literature, we entertain a handful of macro economic variables, including many supposedly forward looking variables: consumption (C); output (Y); inflation (π); the real interest rate (R); and business confidence (BCI). C and Y are measured in log levels, π is measured as yearly changes, while R and BCI are measured in levels

[27] Total factor productivity is not an observable variable, and has to be estimated. As argued in Beaudry and Portier (2006), it may be the case that in response to a technological innovation, properly measured productivity does not increase for a substantial period of time, but that mis-measured productivity responds rapidly due to changes in factor utilization. Our results are robust to using output per hours worked as a measure of (labor) productivity.

(percent). A full description of the variables, their sources and construction is given in Appendix A. Finally, we estimate all model specifications using Gibbs simulations. Details about the estimation procedure are given in Appendix G.3. Here we note that we restrict the model to be stationary when drawing from the conditional posterior. This is done to ensure that the VAR is invertible.

In our baseline specification, we include in the VAR; TFP^a, NI, $OSEBX$, C, π, and R, in that order, and use the Cholesky decomposition to identify the structural shocks. In line with the preceding discussion we will focus on the effects of news and noise shocks.[28]

In the first two columns of Figure 5 we present the results for two versions of the baseline model: One where we use TFP^a to measure total factor productivity (Benchmark model), and one where we use (non adjusted) TFP (Alternative model). We also estimate two versions of the Benchmark model where we exchange consumption (C) with employment (E) and hours worked (H), respectively. The results for these models are reported in the last column of Figure 5 (together with the associated results from the Benchmark). We first discuss the Benchmark results. Following a news shock, productivity and consumption are permanently positively affected. Inflation falls, and stays low for up to 8 quarters, while the real interest rate increases (with some delay). As expected, asset prices also increase, see Figure 14 in Appendix F. Following a noise shock, the initial responses in productivity, consumption, inflation, and the real interest rate, are close to, but not fully in line with the responses following a news shock, but then soon depart. The responses in productivity and the news index (see Figure 14) are not significantly different from zero. Conversely, consumption increase sharply before the effect becomes insignificant after roughly 10 quarters. In line with this, inflation and the real interest rate also increase temporarily. These are interesting results and deserve further comment.

[28] We do not label the unexpected innovations associated with the other variables in the system. Nonetheless, while we tried changing their order it had basically no effect on how news and noise shocks are identified and transmitted though the system. We have also estimated the Benchmark model using a combination of short- and long-run restrictions, where we, as in Beaudry and Portier (2006), restrict consumption to have no long-run effect on itself or productivity. Our main results remain unchanged. These additional results can be obtained on request.

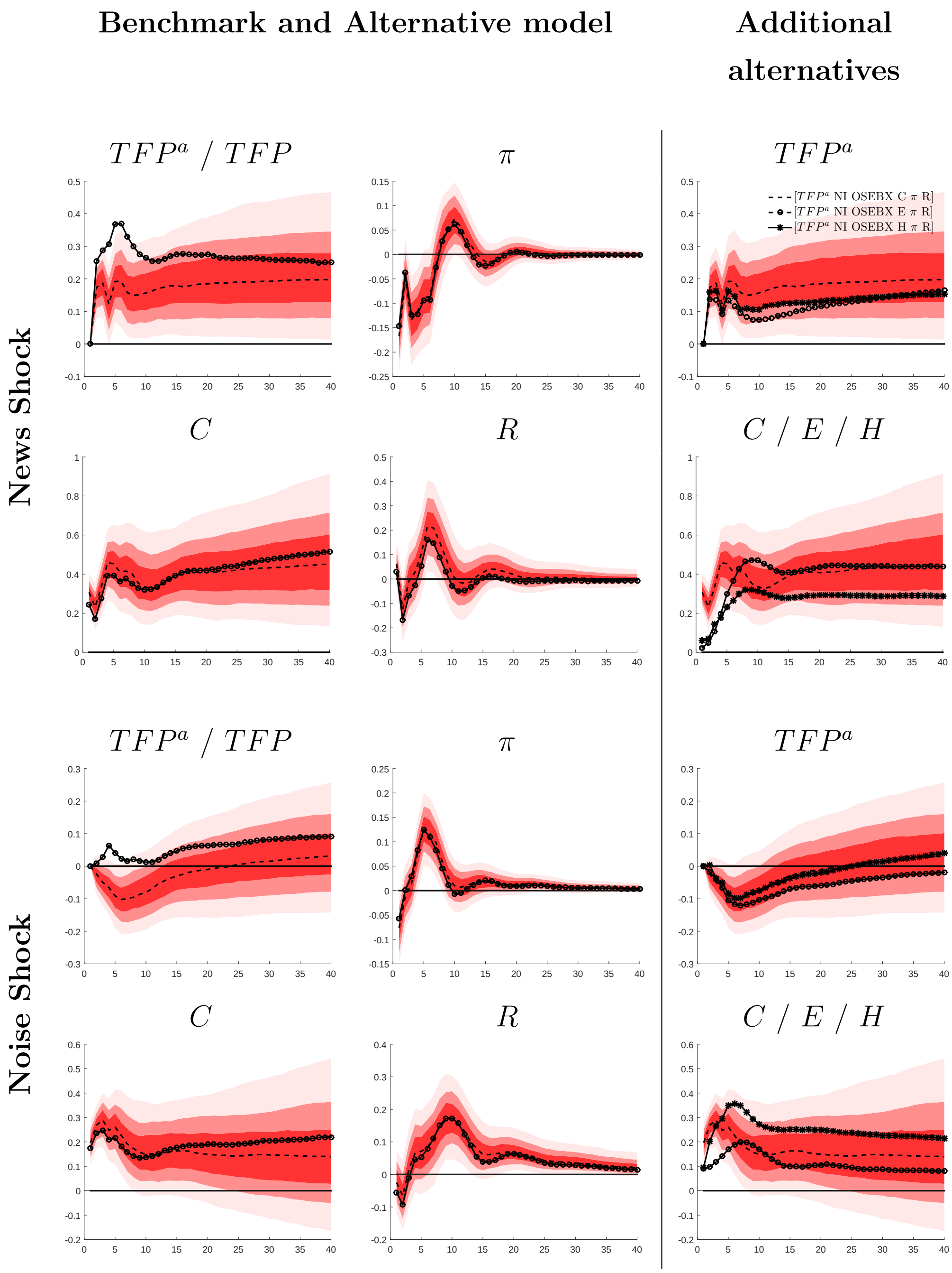


Figure 5. Each graph reports the response (in percent) to an initial one standard deviation shock across response horizons. The color shadings represent the 70, 50, and 30 percent quantiles of the posterior distribution for the Benchmark model. The black dotted line is the associated median estimate. TFP / TFP^a, C, and R are reported in levels. Additional results for asset prices ($OSEBX$) and the news index (NI) are reported in Figure 14 in Appendix F. For π, $OSEBX$ and the NI the responses are reported as yearly growth rates. In the two first columns the black solid and marked line is the median estimate from the Alternative model. In the last column the different alternatives are described by the legend in the upper right corner. E and H are reported in levels.

First, the effects of the identified news and noise shocks are close to those obtained in prominent theoretical news driven business cycle models which include different frictions (like nominal rigidities, adjustment costs, and habit formation), and the agents face a signal extraction problem, see, e.g., Barsky and Sims (2012) and Blanchard et al. (2013). Here, as also exemplified by the simple model in Section 3.1, the news shock is assumed to affect productivity directly, and exerting a permanent effect on the economy, in line with a classical supply shock. The noise shock, on the other hand, is assumed to contain only transitory effects, in line with a typical demand shock. Accordingly, and as documented in Figure 5, after a positive news innovation, consumption and productivity should increase permanently, the real interest rate should rise due to the expected growth in consumption, while inflation should fall. Following a noise shock, consumption and inflation should only grow temporarily, and productivity should be unaffected, since this shock is not associated with changes in economic fundamentals. Again, this is what we observe. Finally, the fact that the noise shock as no significant effect on the news index, suggests that our identification method seems to be able to separate these two components (news and noise) from each other.[29]

Table 1 reports the variance decompositions obtained from the Benchmark (Alternative) model. The news shock explains a non-negligible share of the variation in productivity, consumption, inflation, and asset prices. Depending on which measure we use for productivity, roughly 10 percent of the long-run variation in productivity is explained by news shocks. For asset prices, consumption, inflation, and the real interest rate the news shocks explain around 39, 13, 12, and 7 percent, respectively, of the variability in these variables at the 40 quarter horizon. We also see from the fourth column of the table that the news indicator itself is more or less exogenous to any other disturbances in the model, at least in the short-run. The noise shock, on the other hand, is generally of less importance. Still, for asset prices noise shocks are highly

[29] Comparing the Benchmark results to the Alternative, only for the responses in productivity itself do we observe significant differences. Not surprisingly, the increase in productivity following a news shock is much stronger when we use TFP relative to the TFP^a measure. In the long-run, however, the responses converge. Following a noise shock, productivity is initially more or less unaffected. It then increase slightly, but is never significantly different from zero.

Table 1. Variance decompositions: The numbers reported (in parenthesis) are based on the median impulse response functions from the Benchmark (Alternative) model. See Figure 5.

		Variable							
Shock	**Horizon**	TFP^a	**NI**	**Osebx**	**C**	π	**R**	**E**	**H**
News	1	0.00 (0.00)	1.00 (0.95)	0.50 (0.41)	0.10 (0.07)	0.06 (0.05)	0.01 (0.00)	0.01	0.01
	20	0.07 (0.13)	0.76 (0.63)	0.39 (0.30)	0.17 (0.13)	0.12 (0.11)	0.08 (0.06)	0.20	0.08
	40	0.07 (0.11)	0.73 (0.61)	0.39 (0.29)	0.13 (0.12)	0.12 (0.11)	0.07 (0.06)	0.20	0.08
Noise	1	0.00 (0.00)	0.00 (0.00)	0.48 (0.48)	0.04 (0.03)	0.01 (0.01)	0.00 (0.01)	0.04	0.02
	20	0.02 (0.02)	0.04 (0.04)	0.31 (0.31)	0.05 (0.04)	0.07 (0.06)	0.06 (0.06)	0.04	0.08
	40	0.02 (0.02)	0.04 (0.04)	0.30 (0.30)	0.03 (0.03)	0.07 (0.07)	0.06 (0.06)	0.03	0.06

important, accounting for 48 and 30 percent of the short- and long-run variation, respectively.

For both news and noise shocks the numbers reported in Table 1 are somewhat lower than those typically obtained in the theoretical literature, see, e.g., Lorenzoni (2009), Barsky and Sims (2012), and Blanchard et al. (2013). The numbers are also lower than what's found in recent empirical papers using U.S. data, see, e.g., Barsky and Sims (2011) for news shocks, and Forni et al. (2014) for both news and noise shocks. We do not find this to be particularly surprising for two reasons. First, besides the obvious observation that the identified shocks might differ across studies, Norway, in contrast to the U.S., is a small and open economy, for which a large literature has shown that international business cycle fluctuations matter a lot. According to findings in, e.g., Bjørnland and Thorsrud (2015), international shocks can explain up to 50 percent of the domestic business cycle fluctuations in Norway. If the newspaper data we use is biased toward domestic developments, or the agents in the economy down weight information contained in news topics associated with international developments, such news might be overlooked and therefore lower the variance explained by the news shocks identified here. Second, in some of the above mentioned studies there is only one shock affecting productivity, namely the anticipated news shock. In our model, we have both unanticipated productivity shocks and news shocks. Since our focus has been on identifying news and noise shocks, the results for the unanticipated productivity shocks are not reported (but can be obtained on request). We note, however, that together anticipated and unanticipated productivity shocks account for roughly 60 and 50 percent of the long-run variation in productivity and consumption, respectively.

The last column in Figure 5 reports the responses estimated for the additional model specifications. Following positive news shocks, employment and hours increase almost on impact and are permanently affected. In contrast, following noise shocks the increase is only temporary and becomes insignificant at longer horizons, in line with the response in consumption. The responses for inflation and the real interest rate remain almost identical to those obtained using the Benchmark model. For brevity we do not report them. Finally, the last two columns of Table 1 confirm that the news shock is far from unimportant in explaining the variation in employment and hours worked, with up to 20 percent of the variation in employment explained by this shock.[30]

3.4 What is the news?

Equation (6), in Section 3.1, tells us that news shocks are new information about future productivity. But, what type of information is this exactly? The theory model is embarrassingly silent about such a question. So are bivariate (or larger systems) VARs using asset prices (or confidence measures) to identify news shocks. In the latter case all we typically know is that news shocks are orthogonal to current productivity innovations. However, such shocks can be a linear combination of many potential news items. Using our suggested news index we can come closer to an answer. Since the aggregated news index is a weighted average of all the individual news topics, the structural news shocks would be as well. This allows us to investigate what type of news actually constitute a news shock by looking at the history of structural shocks, and how

[30] We have also estimated the Benchmark model exchanging consumption (C) with investment (I) and output (Y), respectively. For readability the results from these experiments are not reported, but we observe that a news shock cause an increase in investment followed by a boom lasting for roughly 2 years. After 3 years the response converge back to zero. Following a noise shock the response in investment is insignificant at all horizons. For output, the results resemble those for consumption, i.e., a news shock leads to a permanently higher output level while a noise shock causes only a temporary boom.

these are decomposed into different news topics.

In the upper part of Figure 6 the news shock decomposition is illustrated using color shadings. At each point in time the height of the bar reflects the size of the structural news shock while the color shadings are the contribution of each individual news topic, scaled by their weight and likelihood of being in the corpus in that quarter. Technically, these historical shocks are computed based on the Benchmark VAR described in the previous section. We note, however, that irrespective of which of the VARs reported in Figure 5 we use, or the simpler bivariate VARs estimated in Appendix D, the results are close to identical.[31] We start by looking at the overall picture, where some well-known periods stand out. In the aftermath of the financial crisis and during the Great Recession we observe a series of negative news shocks affecting the Norwegian economy. The biggest negative news shock during this period hit the Norwegian economy early in 2008. A series of positive and negative news shocks are also clearly visible for the late 1990s and early 2000s. These are periods associated with high growth in the Norwegian economy (late 1990s), the Asian financial crisis (around 1998), and the dot-com-led recession in the U.S. (early 2000s).

As indicated by the color shadings in Figure 6, many news topics contribute to the news shocks across time. The lower part of Figure 6 reports the average contribution of each topic across the sample. For readability we only report the topics that cumulatively explain up to 60 percent of the shock. The most important topic by far is the *Funding* (14) topic, which on average contributes roughly 25 percent to the structural news shock in each period. Thereafter follows general topics as *Support* (10), *Stock market* (12), *Monetary policy* (35), *IT/Startup* (33), *Results* (68), and *Fear* (77), but also topics especially relevant for the Norwegian economy, e.g., *Oil price* (55) and *Oil production* (44).[32]

Through which channels do news shocks affect productivity? As documented in Figure 5, news shocks cause a large and persistent increase in productivity, and explain a non-negligible share of its long-run variation. Consistent with

[31]The correlation between the historical news shocks across models is as high as 0.98 and never below 0.95. This is so because the news index is more or less exogenous, rendering which variables to include in the SVAR and their ordering of less importance for how we identify the news shock.

[32]Still, across time, we also observe significant variation. The discussion of this variation is for brevity delegated to Appendix E, where we focus on two specific episodes.

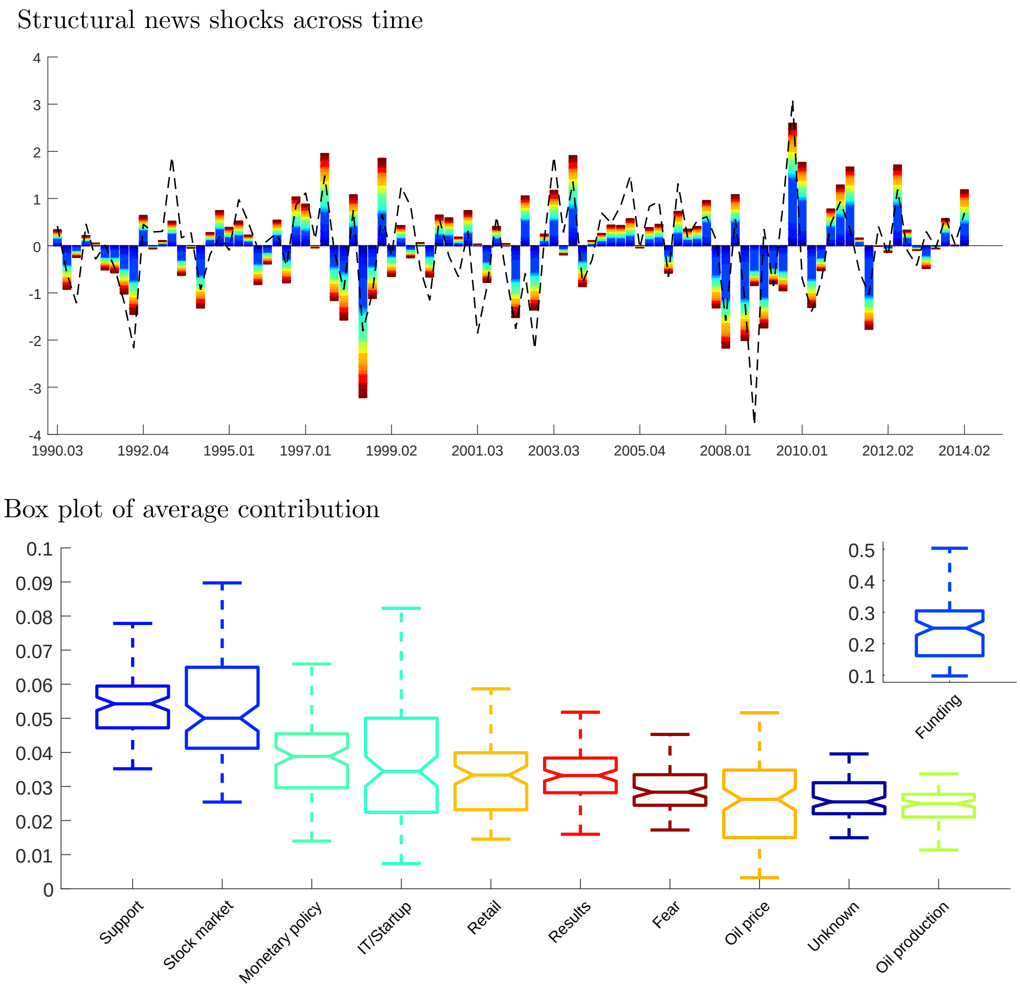


Figure 6. The hight of the bars in the first graph reflect the total contribution from each of the news topics, scaled by their weight and likelihood of being in the corpus at every time period. Each color segment is associated with one particular topic. The black dotted line is the historical news shocks implied by the $[TFP^a\ OSEBX]$ model from Figure 9 in Appendix D. The second graph reports a simple box plot of the contribution (in percent) of each topic to the aggregate news shock across the sample. For each box, the central mark is the median, the edges of the box are the 25th and 75th percentiles, and the whiskers extend to the most extreme data points not considered outliers.

this, it is easy to argue that topics as *IT/Startup*, *Oil price*, and *Oil production* represent news that likely moves the production frontier of the economy outward and thereby increases the productive potential. Coupled with the fact that productivity increases almost on impact supports a classical interpretation of news shocks where they contain new information about future productivity directly. In contrast, the *Monetary policy*, *Results*, *Fear*, and, not least, the *Funding* topics are also found to be important, although none of them is typically associated with productivity developments directly. But, as shown in many recent papers, unexpected shocks to borrowing conditions can, on their own, generate a large and persistent recession with a drop in both output, consumption and productivity, see, e.g., Khan and Thomas (2013) and Miao and Wang (2012). In relation to this, the future level of the interest rate is naturally important for borrowing conditions, making news about *Monetary policy* potentially relevant to future productivity. Finally, the importance of uncertainty, or fear, in explaining large economic fluctuations has been highlighted in a number of studies following the financial crisis, with perhaps the seminal paper by Bloom (2009) being the most well known.[33] However, whether unexpected changes in uncertainty on its own lead to transitory or permanent effects on productivity and output is debated, see, e.g., the discussion in Bloom (2014). Our results indicate that they have permanent effects.

4 Implications and additional results

The empirical findings documented above have two important implications. First, our results indicate that models where innovations in asset prices are used as a proxy for news shocks most likely confound the effects of news and noise shocks. This can be seen clearly from the results reported in Figure 5 and Table 1. In the short-run, news and noise shocks explain almost all of the variation in $OSEBX$, indicating that movements in returns are well explained by these two shocks alone. At the same time, as shown in Figure 5, the effects of news and noise shocks on the general economy are very different. The same finding can also be obtained from Figure 6 where we plotted the time series history of the structural news shocks derived from our Benchmark model. In the figure, we also reported the history of news shocks as implied by one of the models which uses innovations to asset prices as a proxy for news shocks, see Appendix D. Comparing the colored bars with the dotted black line in the figure, we observe a substantial difference between the two. The correlation is 0.66. However, if asset prices contain both news and noise, as implied by

[33] Interestingly, the correlation between the *Fear* topic, see Figure 2, and the U.S. VIX index, which is an often-used proxy for economic uncertainty, is well above 0.7. For a related measure using Norwegian data, computed by the authors based on quarterly standard deviations of asset returns, the correlation is just below 0.7.

our Benchmark model identification scheme, one would think the correlation between news shocks identified using asset prices as a proxy, and the combined effect of news and noise from our Benchmark model, would yield a correlation closer to unity. This intuition is correct. The correlation is 0.94.

Second, regarding the exact interpretation of news shocks, there is little consensus. Two opposing views, based on empirical evidence, are reflected in the influential papers by Beaudry and Portier (2006) and Barsky and Sims (2011). Our interpretation of a news shock differs from both. In Barsky and Sims (2011), but in contrast to what we show, news shocks cause a negative co-movement among productivity, and output and hours worked. As argued in Beaudry and Portier (2014), this suggests that the effect of news shocks may actually be to create a recession, as would be consistent with a Real Business Cycle (RBC) model, as opposed to creating a boom. Still, our interpretation of a news shock does not accommodate the (contrasting) interpretation pursued following Beaudry and Portier (2006) either. Here, news shocks about future productivity can set off a boom today, while a realization of productivity which is worse than expected can induce a bust without any actual reduction in productivity itself ever occurring.

However, our main results do resemble those obtained in prominent news driven business cycle models where news is restricted to affecting future productivity directly, see, e.g., Barsky and Sims (2012) and Blanchard et al. (2013), but also the simple model outlined in Section 3.1. In line with this, as a final robustness experiment we have re-estimated the Benchmark model with an aggregated news index measure based on how well the news topics predict future TFP, confer the discussion in Section 2.3. The results from this experiment are reported in Figure 15 in Appendix F. In essence, the effects of a news shock are very close to those reported in Figure 5. As such, our empirical experiment, although highly data driven, seems to confirm key theoretical predictions.[34]

On the other hand, the result that a broad range of news topics actually contributes significantly to news shocks, see Figure 6 and the discussion in Section 3.4, questions the validity of the standard assumption about how news is supposed to affect productivity, but suggests that it's not a bad approxima-

[34] Our results are also robust to the assumption that productivity is contemporaneously endogenous, i.e., allowing news shocks to affect productivity contemporaneously, see Figure 15 in Appendix F.

tion. Alternatively, the theory models might give the correct predictions, but for the wrong reasons. To be concrete, in light of our decomposition results, the dynamic process for productivity given by equation (6) seems somewhat simplistic. Is there just one ϵ_{t-1}? Likely not, and there are potentially no good reasons to believe that news shocks about credit and borrowing conditions have exactly the same propagation mechanism as news shocks about the energy sector. However, if this is true, and since we use an aggregate news index, one can easily criticize the identification scheme used in this paper as well. We are fully sympathetic to this objection. Still, the same critique can then be made of all papers that use other news proxies to measure the effect of news shocks. As long as we do not know what the news is about, we cannot know anything about the channels through which it most likely operates either. However, this paper casts light on precisely this conundrum. An interesting area for future research is therefore to investigate the potential heterogeneity in economic responses to different types of news shocks, for example using the methodology proposed in this paper. Likewise, our findings should be suggestive for future theoretical work on how news shocks transmit and ultimately affect productivity and economic fluctuations.

5 Conclusion

The main motivation for this paper has been to construct a more direct measure of news, and evaluate its usefulness in explaining economic fluctuations in line with the news driven business cycle view. The finding that an LDA decomposition of the biggest business newspaper in Norway yields news topics which are easy to interpret and have marginal predictive power for many important economic aggregates, including asset prices, confirms the hypothesis we started with: the more intensive a given topic is represented in the newspaper at a given point in time, the more likely it is that this topic represents something of importance to the economy's future needs and developments. Moreover, using our suggested aggregated news index in a SVAR analysis yields predictions which are consistent with theory models where the agents face a signal extraction problem. Following news shocks inflation fall, the real interest rate rise, while output, consumption, employment, hours and TFP increase persistently.

Following noise shocks output, consumption, employment, and inflation rise for a short time period, only to fall back again. Interestingly, the most important news topic contributing to the news shocks by far is related to developments in the financial markets, credit and borrowing; but many other topics make significant contributions. Among these, and especially important in the Norwegian economy, are topics associated with the energy sector.

The empirical findings documented in this paper have two important implications. First, our results indicate that models where innovations in asset prices are used as a proxy for news shocks most likely confound the effects of news and noise shocks. Second, the decomposition of the news shock into news topics should be suggestive for future work on how news shocks theoretically transmit and ultimately affect productivity and economic fluctuations.

References

Angeletos, G.-M. and J. La'O (2013). Sentiments. *Econometrica 81*(2), 739–779.

Arezki, R., V. A. Ramey, and L. Sheng (2015). News Shocks in Open Economies: Evidence from Giant Oil Discoveries. NBER Working Papers 20857, National Bureau of Economic Research, Inc.

Barsky, R. B. and E. R. Sims (2011). News shocks and business cycles. *Journal of Monetary Economics 58*(3), 273–289.

Barsky, R. B. and E. R. Sims (2012). Information, Animal Spirits, and the Meaning of Innovations in Consumer Confidence. *American Economic Review 102*(4), 1343–77.

Beaudry, P., D. Nam, and J. Wang (2011). Do Mood Swings Drive Business Cycles and is it Rational? NBER Working Papers 17651, National Bureau of Economic Research, Inc.

Beaudry, P. and F. Portier (2006). Stock Prices, News, and Economic Fluctuations. *American Economic Review 96*(4), 1293–1307.

Beaudry, P. and F. Portier (2014). News-Driven Business Cycles: Insights and Challenges. *Journal of Economic Literature 52*(4), 993–1074.

Bjørnland, H. C. and L. A. Thorsrud (2015). Boom or gloom? Examining the Dutch disease in two-speed economies. *Economic Journal (forthcoming).*

Blanchard, O. J., J.-P. L'Huillier, and G. Lorenzoni (2013). News, Noise, and Fluctuations: An Empirical Exploration. *American Economic Review 103*(7), 3045–70.

Blei, D. M. (2012). Probabilistic topic models. *Communications of the ACM 55*, 77–84.

Blei, D. M. and J. D. Lafferty (2006). Dynamic topic models. In *Proceedings of the 23rd International Conference on Machine Learning*, ICML '06, New York, NY, USA, pp. 113–120. ACM.

Blei, D. M., A. Y. Ng, and M. I. Jordan (2003). Latent Dirichlet Allocation. *J. Mach. Learn. Res. 3*, 993–1022.

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), 153–76.

Boudoukh, J., R. Feldman, S. Kogan, and M. Richardson (2013). Which News Moves Stock Prices? A Textual Analysis. NBER Working Papers 18725, National Bureau of Economic Research, Inc.

Chang, J., S. Gerrish, C. Wang, J. L. Boyd-graber, and D. M. Blei (2009). Reading tea leaves: How humans interpret topic models. In Y. Bengio, D. Schuurmans, J. Lafferty, C. Williams, and A. Culotta (Eds.), *Advances in Neural Information Processing Systems 22*, pp. 288–296. Curran Associates, Inc.

Dominguez, K. M. E. and M. D. Shapiro (2013). Forecasting the recovery from the great recession: Is this time different? *American Economic Re-*

view 103(3), 147–52.

Forni, M., L. Gambetti, M. Lippi, and L. Sala (2014). Noisy News in Business Cycles. Working Papers 531, IGIER (Innocenzo Gasparini Institute for Economic Research), Bocconi University.

Griffiths, T. L. and M. Steyvers (2004). Finding scientific topics. *Proceedings of the National academy of Sciences of the United States of America 101*(Suppl 1), 5228–5235.

Gutierrez, R. C. and E. K. Kelley (2008). The long-lasting momentum in weekly returns. *The Journal of Finance 63*(1), 415–447.

Hansen, S., M. McMahon, and A. Prat (2014). Transparency and Deliberation within the FOMC: A Computational Linguistics Approach. CEP Discussion Papers 1276, Centre for Economic Performance, LSE.

Kass, R. E. and A. E. Raftery (1995). Bayes factors. *Journal of the American Statistical Association 90*(430), 773–795.

Keynes, J. (1936). *The General Theory of Employment, Interest and Money.* London: MacMillian.

Khan, A. and J. K. Thomas (2013). Credit Shocks and Aggregate Fluctuations in an Economy with Production Heterogeneity. *Journal of Political Economy 121*(6), 1055 – 1107.

Koop, G. and D. Korobilis (2010). Bayesian Multivariate Time Series Methods for Empirical Macroeconomics. *Foundations and Trends(R) in Econometrics 3*(4), 267–358.

Lorenzoni, G. (2009). A Theory of Demand Shocks. *American Economic Review 99*(5), 2050–84.

Mcauliffe, J. D. and D. M. Blei (2008). Supervised topic models. In J. Platt, D. Koller, Y. Singer, and S. Roweis (Eds.), *Advances in Neural Information Processing Systems 20*, pp. 121–128. Curran Associates, Inc.

Mertens, K. and M. O. Ravn (2012). Empirical evidence on the aggregate effects of anticipated and unanticipated US tax policy shocks. *American Economic Journal: Economic Policy 4*(2), 145–81.

Miao, J. and P. Wang (2012). Bubbles and total factor productivity. *American Economic Review 102*(3), 82–87.

Nakajima, J. and M. West (2013). Bayesian Analysis of Latent Threshold Dynamic Models. *Journal of Business & Economic Statistics 31*(2), 151–164.

Pigou, A. (1927). *Industrial Fluctuations.* London: MacMillian.

Ramey, V. A. and M. D. Shapiro (1998). Costly capital reallocation and the effects of government spending. *Carnegie-Rochester Conference Series on Public Policy 48*(1), 145–194.

Romer, C. D. and D. H. Romer (2010). The macroeconomic effects of tax changes: Estimates based on a new measure of fiscal shocks. *American Economic Review 100*(3), 763–801.

Schmitt-Grohe, S. and M. Uribe (2012). What's news in business cycles. *Econo-*

metrica 80(6), 2733–2764.

Sims, C. A. and T. Zha (2006). Does Monetary Policy Generate Recessions? *Macroeconomic Dynamics 10*(02), 231–272.

Soo, C. K. (2013). Quantifying Animal Spirits: News Media and Sentiment in the Housing Market. Working Paper 1200, Ross School of Business, University of Michigan.

Stock, J. H. and M. W. Watson (2003). Forecasting output and inflation: The role of asset prices. *Journal of Economic Literature 41*(3), 788–829.

Tetlock, P. C. (2007). Giving content to investor sentiment: The role of media in the stock market. *The Journal of Finance 62*(3), 1139–1168.

Tetlock, P. C., M. Saar-Tsechansky, and S. Macskassy (2008). More Than Words: Quantifying Language to Measure Firms' Fundamentals. *Journal of Finance 63*(3), 1437–1467.

Vega, C. (2006). Stock price reaction to public and private information. *Journal of Financial Economics 82*(1), 103 – 133.

Watson, M. W. (1986). Vector autoregressions and cointegration. In R. F. Engle and D. McFadden (Eds.), *Handbook of Econometrics*, Volume 4 of *Handbook of Econometrics*, Chapter 47, pp. 2843–2915. Elsevier.

Appendices

For Online Publication

Appendix A Data

The news index (NI) is constructed based on the DN corpus, an LDA decomposition, and predictive results. See Sections 1 and 2.3 of the main paper for details.

Consumption (C), output (Y), investment (I), and employment (E) were obtained from Statistics Norway (SSB). Consumption is defined as "Final consumption expenditure of households and NPISHs"; output is defined as "Gross domestic product, market value"; investment is defined as "Gross fixed capital formation"; and employment is defined as "Employed persons. Employees and self-employed." Consumption, investment, and output are all measured in constant 2012 prices (million NOK). Employment is measured in 1000 persons. All series are seasonally adjusted by the original source.

The business confidence indicator (BCI) and inflation (π) were obtained from the FRED database maintained by the Federal Reserve Bank of St. Louis. A pure consumer confidence indicator, covering the sample needed here, does not exist for Norway. The BCI indicator is a seasonally adjusted confidence index for the manufacturing sector, whose normal value is 100. The original source is OECD, "Main Economic Indicators - complete database." The inflation series is the consumer price index for all items, normalized to 100 in 2010. We seasonally adjust this series using the X12-ARIMA package developed by the U.S. Census Bureau. The real interest rate (R) is constructed using the 3-month interbank rate for Norway (NIBOR), deflated using the inflation series described above. The nominal interest rate was collected from the FRED database maintained by the Federal Reserve Bank of St. Louis.

Asset prices ($OSEBX$) were obtained from Yahoo finance. The $OSEBX$ index is the Oslo Stock Exchange Benchmark Index. As of 2013 it consists of roughly 55 companies listed on the Oslo Stock Exchange.

Finally, the total factor productivity index is constructed based on a simple growth accounting framework, using output (Y), hours worked (H), and the capital stock (KS). For many countries, official statistics exist for total factor productivity. For Norway, none exist with a quarterly frequency. We construct two measures, one as $TFP = log(Y_t/H^{\alpha}(KS_t)^{1-\alpha})$, and the other as $TFP^a = log(Y_t/H^{\alpha}(CU_tKS_t)^{1-\alpha})$. For both measures we set $\alpha = 0.33$. TFP^a is adjusted for capacity utilization (CU), while the TFP measure is not. Norges Bank kindly provided us with numbers for the capital stock and hours worked covering the sample needed in this paper. Capacity utilization (CU) was collected from SSB's Business tendency survey for manufacturing, mining and quarrying.

Appendix B Labeling the topics

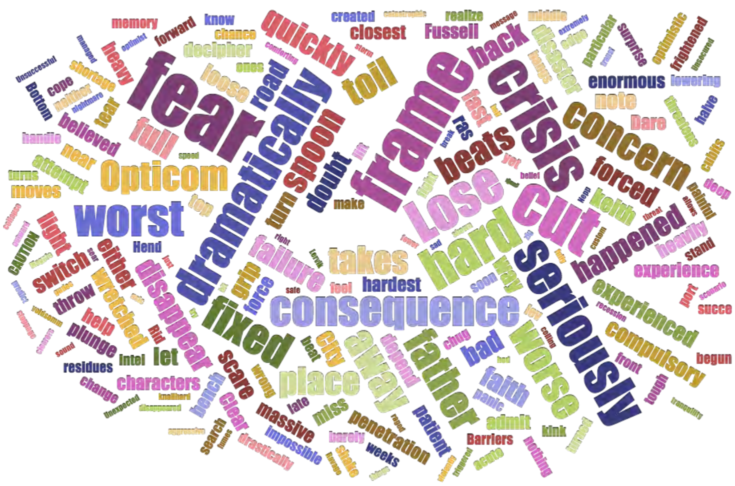
As discussed in Section 1.1 of the main paper, the LDA algorithm does not label the extracted topics. Instead, the outputs from the procedure are probability distributions over all the words in the corpus, and over a set of a predetermined number of topics. One useful way to represent these probability distributions is by using word clouds, where the size of a word within a given topic is drawn to a size that corresponds to the probability of that word occurring in that topic. From this we can then classify the different topics and give them distinct labels. Inherently the latter part of this procedure is based on discretion, but as emphasized in Section 1.1, the exact labeling does not play a material role for the main results of the paper.

Figure 7 exemplifies our procedure for 2 out of the 80 distinct topics. As seen in the figure, for topic 14, the words “loans”, “guarantee”, “interest”, “financial”, and “maturity” are all given a large weight in this topic. Thus, instead of referring to this as topic 14, we classify the topic as a *Funding* topic. To us, this seems to encompass the most important words within this topic reasonably well. Similarly, for topic 77, we end up with the classification *Fear*.

Table 2 reports the full list of estimated topics, labels, as well as a list of the most important words associated with each individual topic. The table also reports the frequency of articles for each topic that are best described by that topic. For most of the estimated topics we find this classification scheme easy to apply. Still, for two topics we found labeling particularly difficult, namely topics 2, and 74.

We note that some words might get lost in translation. That is, the raw data are from a Norwegian newspaper, so the language we work with is Norwegian. However, to reach a wider audience, all words are translated into English. For simplicity, we used Google Translate. One problem with this approach is that the translation of many Norwegian words produces two English words. One example is the Norwegian word “sentralbank”, which translates as “central bank.” The translation here is clear, but since one word is turned into two, “sentralbank” shows up as two distinct elements, “central” and “bank.”

Topic 14: Funding



Topic 77: Fear

Figure 7. Word clouds and topic categorization. For each word cloud the size of a word reflects the probability of this word occurring in the topic. A word cloud is created based on the 200 first words in each topic.

Table 2. Estimated topics and labeling. The topics are labeled based on the meaning of the most important words, see the text for details. The "# of articles" column reports the number of articles, in the full sample, which according to the model, belong to that specific topic. The words are translated from Norwegian to English using Google Translate.

Topic	Label	# of articles	First words
Topic 0	Anglo-Saxon	4457	the, new, of, york, doll, and, in, london, world, street, is, you, on, english, england, wall
Topic 1	Leadership	4357	position, forests, chairman, president, ceo, dismissal, executive, candidate, elected
Topic 2	Unknown	5717	smile, night, man, wall, house, door, gate, clock, home, minute, no, night, black
Topic 3	Knowledge	707	know, things, think, answer, never, good, feel, always, really, need, tell, pretty, just, feel, try
Topic 4	Context	710	degree, power, unequal, change, influence, context, difference, high, impact, significantly
Topic 5	Public safety	7497	police, finance guards, sight, illegal, investigation, indicted, prison, corruption, report
Topic 6	Government policy	3948	suggestions, parliamentary, department, selection, treasury, minister, change, foss, budget
Topic 7	Olympics	2301	olymics, participate, visit, invite, lillehammer, interest, business, gold, arranging, walk, story
Topic 8	Cooperation	1435	group, cooperation, establish, trap, tandberg, strategy, data, ulltveit, develop, abb, alliances
Topic 9	Manufacturing	6850	product, production, produce, factory, manufacturer, brand, bet, competition, marketing
Topic 10	Support	1638	support, establish, organize, funding, culture, advice, help, freely, purposes, create, shape
Topic 11	Sweden	4894	swedish, sweden, nordic, north, stockholm, finland, finnish, ericsson, denmark, ab, island
Topic 12	Stock market	9519	exchange, fell, quotes, steps, investor, stockmarket, index, points, upswing, decreasing
Topic 13	Automobiles	6991	car, model, engine, drive, ford, volvo, toyota, mercedes, bmw, class, saab, brand
Topic 14	Funding	4916	loans, interest, equity, guarantee, funding, finance, financial, bond, risk, financial crisis
Topic 15	Employment benefits	5968	public, private, scheme, sector, pension, payday, measures, labor, working, service

Continued on next page

Table 2 – continued from previous page

Topic	Label	# of articles	First words
Topic 16	Art	7169	picture, art, exhibition, gallery, artists, museum, munch, painting, auction, design
Topic 17	Sport	4815	games, game, club, soccer, sponsoring, sports, world cup, cosmopolitan, skiing, jahr, win
Topic 18	Europe	6082	german, germany, european, french, euro, france, spain, italy, spanish, berlin, italian
Topic 19	IT/Technology	8872	internet, online, technology, pc, microsoft, service, system, electronic, apple, user, machine
Topic 20	Conflict	5978	war, iraq, military, attack, forces, al, conflict, defense, iran, israel, nato, soldier, un, vest
Topic 21	Success	411	top, list, space, good, happy, road, number, eight, seven, loud, close, joy, promise, right
Topic 22	Communication	7701	telenor, mail, mobile, customer, netcom, hermans, telia, online, vimpelcom, telecom
Topic 23	Brokerage firms	4442	customer, brokerage, trading, bonus, trade, securities, brokerage, acta, industry
Topic 24	Reasoning	774	should, therefore, quite, moreover, faith, sure, namely, right, of course, interesting, hardly
Topic 25	Family	4597	woman, children, men, family, young, father, man, home, mother, age, parents, age, son
Topic 26	Food	4913	wine, food, restaurant, taste, salt, nok, pepper, drinks, fruit, fresh, bottle, menu, server
Topic 27	Investigation	1103	report, investigate, analysis, conclusions, assessment, conducted, conclude, answers, base
Topic 28	Shipping	12441	ships, shipping, dollar, wilhelms, fleet, proud, frontliners, berges, tank, rat, skaug, ugland
Topic 29	Criticism	2886	criticism, express, asserting, article, claim, incorrectly, press, pr, react, should, respond
Topic 30	LLC	4120	llc, group, family, dividend, asset, holding, equity, subsidiary, ownership, shareholder
Topic 31	East Asia	8142	china, japan, chinese, asia, japanese, indians, dollar, government, kong, brazil, korea, south
Topic 32	Aviation	8951	sas, fly, travel, airline, english, braathens, airport, passenger, gardemoen, color, air, traffic
Topic 33	IT/Startup	5112	it, group, acquisitions, partner, establish, business, entrepreneur, steen, industry, office
Topic 34	UK/US presidents	5428	british, london, president, uk, election, pound, bush, obama, political, clinton, conservative
Topic 35	Monetary policy	11863	interest, central bank, inflation, point, governor, percentage points, fell, steps, economy

Continued on next page

Table 2 – continued from previous page

Topic	Label	# of articles	First words
Topic 36	Industry	4001	industry, industries, workplace, business, create, small, competition, help, better, develop
Topic 37	Rig	8420	issue, rig, dollar, offshore, collier, shareholder, drilling, retrieve, seadrill, sundal, pareto
Topic 38	Life	4551	human, history, words, live, feel, kind, shape, death, man, old, him, never, express, modern
Topic 39	Newspapers	10603	newspaper, media, press, schibsted, Dagbladet, journalist, vg, eve mail, editor
Topic 40	Negotiations	684	solution, negotiation, agreement, parties, confirm, offers, conversation, process, negotiate
Topic 41	EU	8997	eu, ef, eea, commission, membership, no, brussel, eft, farmers, negotiations, agriculture
Topic 42	TV	10852	television, nrk, channel, advertising, radio, digital, media, agency, program, commercial
Topic 43	Financial supervision	4290	letter, information, financial supervision, enlightenment, auditors, control, accounting
Topic 44	Oil production	10415	statoil, hydro, oil, field, gas, oil company, shelf, stavanger, platform, shell, findings
Topic 45	Charity	3515	south, organization, africa, church, poor, help, congo, red, aid, rich, big, un, trade council
Topic 46	Justice	6206	lawyer, judge, appeals, claims, supreme court, claim, lawsuit, district court, strife, legal
Topic 47	Literature	7430	book, read, books, reading, writing, history, novel, writer, no, name, him, acted, author
Topic 48	Calender	725	week, previous, january, march, monday, friday, october, december, november, february
Topic 49	Aker	7591	aker, kværner, røkke, finance, option, tdn, rgi, shareholder, hafslund, enlighten, cruise
Topic 50	Projects	2353	project, cost, investment, cover, construction, operation, expansion, budget, annual
Topic 51	Nature	4314	water, meter, city, boat, mountains, ocean, outside, accident, weather, human, earth
Topic 52	Denmark	2017	danish, foreign, denmark, norwegians abroad, immigration, copenhagen, outdoors
Topic 53	Fishery	7457	fish, salmon, tons, seafood, food, marine, fishing, pan, fjord, norway, boat, plant, kilo
Topic 54	Retail	9795	shop, hotel, brand, trondheim, hotel, rema, reitan, ica, coop, stordalen, norgesgruppen
Topic 55	Oil price	8118	dollar, oil, barrel, brokerage, first, analyst, opec, analysts, analyst, fell, steps, securities

Continued on next page

Table 2 – continued from previous page

Topic	Label	# of articles	First words
Topic 56	Energy	11634	energy, emissions, tons, industry, statkraft, elkem, production, aluminum, cent
Topic 57	Savings banks	5535	loss, savings, focus, kreditkassen, lost, middle, lending, positive term, bank manager
Topic 58	Expertise	3411	leader, experience, often, organization, create, people, experience, challenge, thinking
Topic 59	Offshore	8418	contract, shipyards, supply, contract, signed, offshore, building, siem, equipment, kongsberg
Topic 60	Institutional investing	5241	fund, investor, investment, returns, investing, risk, managing, capital, place, private oil fund
Topic 61	Russia	6155	russia, west, russian, east, poland, moscow, president, soviet union, ukraine, authorities
Topic 62	Education	6675	school, university, professor, student, educate, research, studies, subjects, bi, institute
Topic 63	Health care	5034	hospital, physician, health, patient, human, treatment, medicine, help, expensive, develop
Topic 64	Shareholders	5456	orkla, shareholder, chairman, competition, general, bankruptcy, creditor, investor
Topic 65	Macroeconomics	8457	economy, unemployment, lower, forecast, economist, consumption, high, demand
Topic 66	Housing	11098	housing, property, real estate, apartment, square, houses, condos, land, rent, move
Topic 67	Government regulations	3360	rules, government, competition, regulations, prohibitions, competition authorities
Topic 68	Results	1545	number, growth, average, proportion, increase, decrease, compare, roughly, city
Topic 69	Publishing	4425	publishing, books, book, gyldendal, cappelen, smith, aschehoug, book club, copy, nygaard
Topic 70	Norwegian politics	11054	party, right, ap, labour, stoltenberg, political, frp, sv, election, parliamentary, politics, left
Topic 71	Norwegian counties	8035	municipality, trondheim, north, tromsø, nsb, county, local, municipal, kristiansand
Topic 72	Taxation	5171	tax, income tax, wealth tax, property, remove, lower paid, amount, system, compute
Topic 73	Quarterly results	9746	quarter, deficit, surplus, operating, tax, third one, half, group, fourth, minus, last year
Topic 74	Unknown	557	him, took, did, never, later, began, stood, gave, name, old, man, did, thought, happened
Topic 75	Entertainment	11182	film, music, record, play, artist, movie, cd, band, singing, playing, public, record, scene

Continued on next page

Table 2 – continued from previous page

Topic	Label	# of articles	First words
Topic 76	Labor unions	8077	lo, nho, members, pay, union, strike, organization, settlement, los, union, valla, settlements
Topic 77	Fear	847	locked, fear, frame, cutting, crisis, hard, seriously, lose, dramatically, worst, consequence
Topic 78	Private banking	5975	dnb, storebrand, north, merger, bud, mutual insurance, uni, insurers, shareholders
Topic 79	Public debate	2617	political, society, debate, power, politics, politician, politicians, public, system, roll

Appendix C Topics versus sentiment

To gauge the extent to which our news topics approach also adds marginal predictive value relative to a sentiment based approach, we ran a set of alternative predictive experiments. In particular, we first constructed a sentiment index based on keywords. We then used this index in the ARX(p) regressions described in Section 2, and finally, we compare the posterior odds ratio between this sentiment augmented model relative to the ARX(p) models using news topics. Below, we describe in greater detail how the sentiment index is constructed and report the predictive results.

The sentiment based index is constructed based on the whole news corpus, see Section 1. From this corpus we simply count the number of positive and negative words, which is the standard methodology employed in this branch of the literature, see, e.g., Tetlock (2007). To define which words are positive/negative we use an external word list, namely the *Harvard IV-4 Psychological Dictionary*. This is a list with English words, containing many synonyms, and translating this list to Norwegian is not trivial. Our word list defining positive and negative words are based on this list, but not a direct translation. The final set of words consist of 40 positive and 39 negative words. The list can be obtained upon request.

The count procedure delivers two time series, covering the sample from May 2 1988 to December 29 2014, each containing the number of positive and negative words in the news each day. These series were then normalized such that each daily observation reflected the fraction of positive and negative words,

Figure 8. The sentiment index, OSEBX and a measure of the business cycle. The business cycle measure is the Hodrick Prescott (HP, $\lambda = 40000$) filtered GDP. All series are standardized.

i.e.:

$$Pos_t = \frac{\#positivewords}{\#totalwords} \quad Neg_t = \frac{\#negativewords}{\#totalwords} \tag{11}$$

Finally, in line with the literature, the sentiment index itself is defined as:

$$S_t = Pos_t - Neg_t \tag{12}$$

Figure 8 reports the estimated sentiment index together with a measure of the Norwegian business cycle and asset prices. As seen from the figure, the sentiment index seems to have some leading properties relative to the business cycle measure. However, most striking is the close resemblance between the sentiment index and the developments in the asset market. Apart from some deviations early on in the sample, the two series track each other closely. That said, the sentiment index does not, on average, seem to lead the asset market. This is particularly evident during the 2000s, and in the early periods of the Great Recession.

The posterior odds ratios between the sentiment based model and the topic based models are summarized in Table 3. For brevity we only report comparative results for model comparisons where the news topics augmented models receive a higher score for two or more outcome variables, as well as a summary statistic for showing how many topic models $2ln(PO_{ij}) > 0$. Commenting on this latter statistic first, we see that for asset prices, for as many as 17 out of the 80 topic augmented ARX(p) models, the posterior odds ratio is favorable. For output, investment and consumption the numbers are substantially lower, and for BCI and productivity somewhat in between. Thus, compared against all topics, the sentiment index seems to be a good predictor. Still, some topics are

Table 3. News topics versus sentiment. Each column row entry in the table reports the relative marginal likelihood ratio between two competing ARX(p) models, one where we include a news topic as an additional predictive variable, and another where we include a sentiment index. We refer to these as the Alternative (i) and the Null (j) models, respectively. The table only reports the relative marginal likelihood ratio for Alternative models where $2ln(PO_{ij}) > 0$ for at least two of the outcome variables, i.e., for outcomes where the Alternative model receives a higher score than the Null for at least two outcome variables. The last row of the table reports a simple sum, i.e., the number of topic augmented models for which $2ln(PO_{ij}) > 0$ for each outcome variable.

Topic	**Variable**					
	BCI	**OSEBX**	**Y**	**I**	**C**	**TFP**
Stock market (12)	-5.52	2.02	-13.78	2.47	-35.55	-1.74
Funding (14)	-6.40	5.80	-6.85	-3.55	-7.76	13.27
Shipping (28)	-2.43	-1.01	3.47	-5.56	-3.26	1.07
Projects (50)	-9.89	-11.82	-0.55	0.18	-5.74	0.01
Oil price (55)	-0.72	0.83	-12.39	-9.70	-2.54	0.21
Macroeconomics (65)	1.90	-5.68	-2.65	1.31	6.04	1.30
Labor unions (76)	-1.99	0.08	-3.49	-3.81	-11.42	4.61
Fear (77)	-2.18	1.46	-41.77	-3.83	0.30	13.49
All	7.00	17.00	3.00	4.00	3.00	9.00

hard to beat, and perform relatively well across many outcome variables. For example, topic 14 outperforms the sentiment augmented model in predicting both asset prices and productivity. Moreover, the posterior odds ratio strongly suggests the topic model is superior, i.e., the score is well above 2 for these outcome variables. Similar results, although for different outcome variables can be found for the other topics listed in Table 3. Naturally, these topics are also the ones that receive the best score relative to simple AR(p) models, see Section 2.2. And, as discussed in Section 2.2, these topics also make sense from an economical point of view.

Thus, not only do the best performing news topics outperform the sentiment based approach by adding marginal predictive power, they also contain more useful information about what the news actually is about. That said, and as indicated by newer studies in finance, see, e.g., Boudoukh et al. (2013), a more refined construction of the sentiment index can maybe change these predictive results. We leave it for further research to address how well the (statistical) LDA methodology works relative to different (rule-based) word count procedures.

Appendix D Results in a bivariate system

The black solid and dotted lines in Figure 9 report the estimated impulse responses and variance decompositions following a news shock from bivariate SVARs, where we use the news index to identify news shocks using a simple recursive ordering of the variables.[35] For clarity, the legends in the figures show the variables included and their order for each of the models. As seen in the figure, irrespective of which productivity measure we use, TFP or TFP^a, productivity increases sharply already one quarter after the news shock. In addition, a news shock lead to a permanent higher productivity level, although this effect is strongest for the specification entertaining the TFP productivity measure. When news shocks are identified using the news index, news explains up to 15 percent of the variation in productivity at the 3 year horizon.

Compared to the results obtained in the hallmark work by Beaudry and Portier (2006), using U.S. data and asset price innovation to identify news shocks, our results are similar. Still, there is a subtle difference, at least when using TFP^a to measure productivity. In Beaudry and Portier (2006), following news shocks, TFP^a growth picks up with a considerable lag. In contrast, our results indicate that productivity growth picks up almost on impact irrespective of which productivity measure we use. To convince the reader that this result is not because we use our suggested news index to identify the news shock, the gray solid and dotted lines in Figure 9 reports the impulse responses and variance decompositions following a news shock from bivariate SVARs, but where we now use unexpected innovation in asset prices to identify news shocks. As seen from those results, following a positive news shock TFP increases sharply already after one quarter and is permanently affected. That said, when we use the capacity adjusted productivity measure, i.e., TFP^a, news shocks identified using asset prices actually lead to only a short lived increase in productivity followed by a permanent fall. This stands in stark contrast to what is found in, e.g., Beaudry and Portier (2006). Moreover, for neither model specifications do

[35] See Section 3.3 of the main paper for details about the data definitions and estimation, and Appendix G.3 for a description of the VAR methodology.

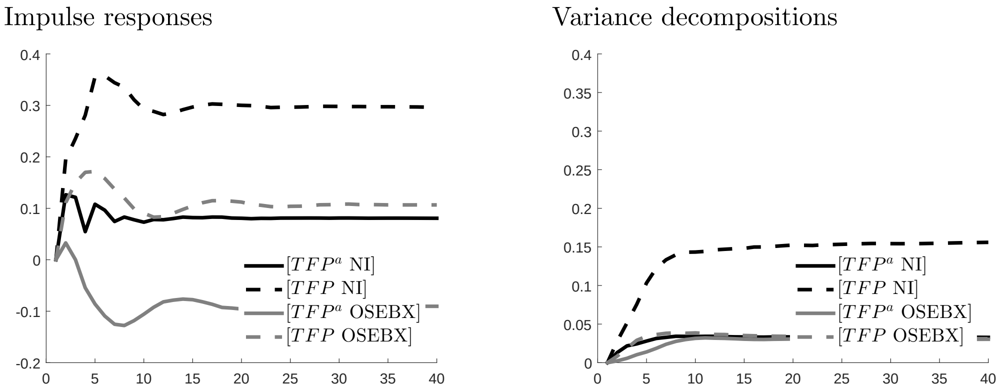


Figure 9. News shocks identified using the news index and stock prices. For all model specifications, the graphs to the left report the response (in percent and levels) of productivity to a one standard deviation shock across response horizons. The graphs to the right report the associated variance decompositions (VDC).

the news shocks explain any significant fraction of the economic fluctuations in productivity.

In sum, and to the extent that news shocks should be associated with future productivity increases (as has been the guiding principle in the news literature), our results show that using the news index to identify a news shock seems to provide more robust results than when using asset prices for the same purpose. As discussed in Sections 4, one potential reason for this is that news shocks identified using innovations to asset prices may be a mixture of news and noise shocks. Another implication of the results described above is that news shocks are not necessarily interpretable as a change in future technological opportunities, as argued in Beaudry and Portier (2006), but maybe more easily interpretable in line with a classical productivity shock, as argued in Barsky and Sims (2012). However, and as mentioned in Beaudry and Portier (2014), these are not the only possible interpretations. For example, an alternative view is to see productivity as endogenous, with short-run non technological intrinsic shocks eventually affecting productivity (with different degrees of delay).

We note that the discrepancies described above are not driven by the combined effect of how we calculate productivity and the usage of the news index either. Barsky and Sims (2012) identify news shocks as unexpected innovations to consumer confidence, and show that these contain incremental information about economic activity and consumption far into the future. Their conclusion

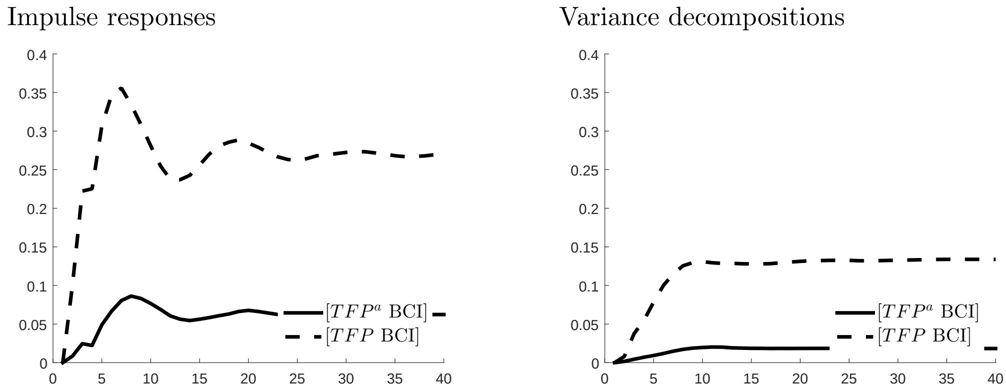


Figure 10. News shocks identified using business confidence. For all model specifications, the graphs to the left report the response (in percent and levels) of productivity to a one standard deviation shock across response horizons. The graphs to the right report the associated variance decompositions (VDC).

strongly suggests that confidence innovations contain true news about the future prospects of the economy. When we estimate the SVARs exchanging asset prices with a confidence measure, as in Barsky and Sims (2012), we obtain very much the same results as when using the news index, see Figure 10.

Appendix E What is the news?

Figure 6 reports the historical news shocks and the average contribution of each news topic across the sample. In the latter figure we also observe significant variation across time. This is documented in Figure 11. For example, in the mid 1990s the Norwegian economy was booming. Among the news topics contributing above average to the positive news shocks in this period were *Stock market* (12), *Retail* (54), *Results* (68), *Offshore* (59), *Energy* (56), *IT/Startup* (33), and *Monetary policy* (35). The first three tells us that this was a general boom in the Norwegian economy. The fact that the *IT/Startup* (33) and *Monetary policy* (35) topics also come across as especially important during this period resonates well with what we today know about this period. It was characterized by many start-ups, and a rapid adoption by private and public institutions of never-seen-before information technology.[36] In the late 1990s,

[36] For example, the number of companies listed on the Oslo Stock Exchange as belonging to the IT industry grew by over 70 percent from the 1980s to the 1990s, which is almost twice as fast as any other industry sector.

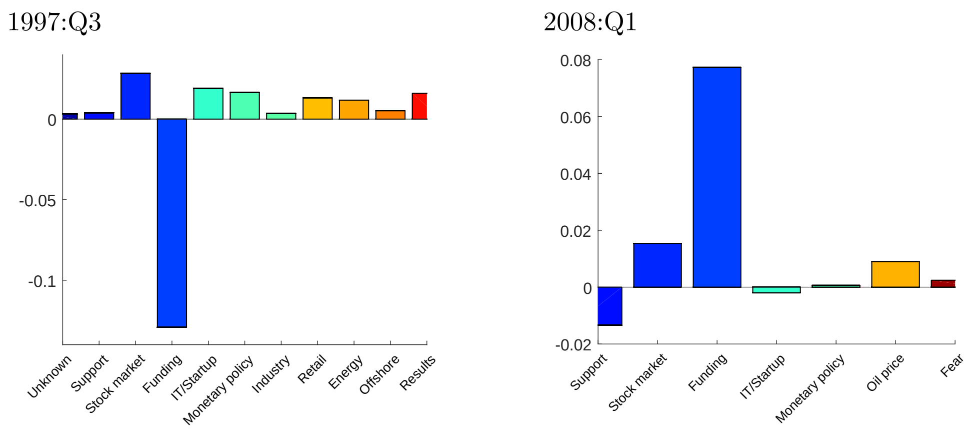


Figure 11. The graphs for the specific time periods report the difference between the average contribution of each topic and that topic's contribution that quarter. See also Figure 6.

the role of the central bank also changed considerable, moving from a quasi-fixed exchange rate regime to inflation targeting, giving monetary policy a more central role in stabilizing economic fluctuations. Turning to a period associated with the Great Recession, i.e., 2008:Q1, we see from Figures 6 and 11 that the large negative news shocks in this quarter are mostly associated with negative news regarding the *Funding* topic, which contributes almost 10 percent more than it does on average. This comes as no surprise; in fact, the phrase *financial crisis* proves as one of the most important for this topic, see Table 2. We also note that news associated with the *Oil price* (55) and *Stock market* (12) topics added more than usual to the aggregate news shock in this quarter, and the news seems more condensed compared to in 1997:Q3, i.e., fewer topics are needed to explain up to 60 percent of the total shock.

Appendix F Additional results

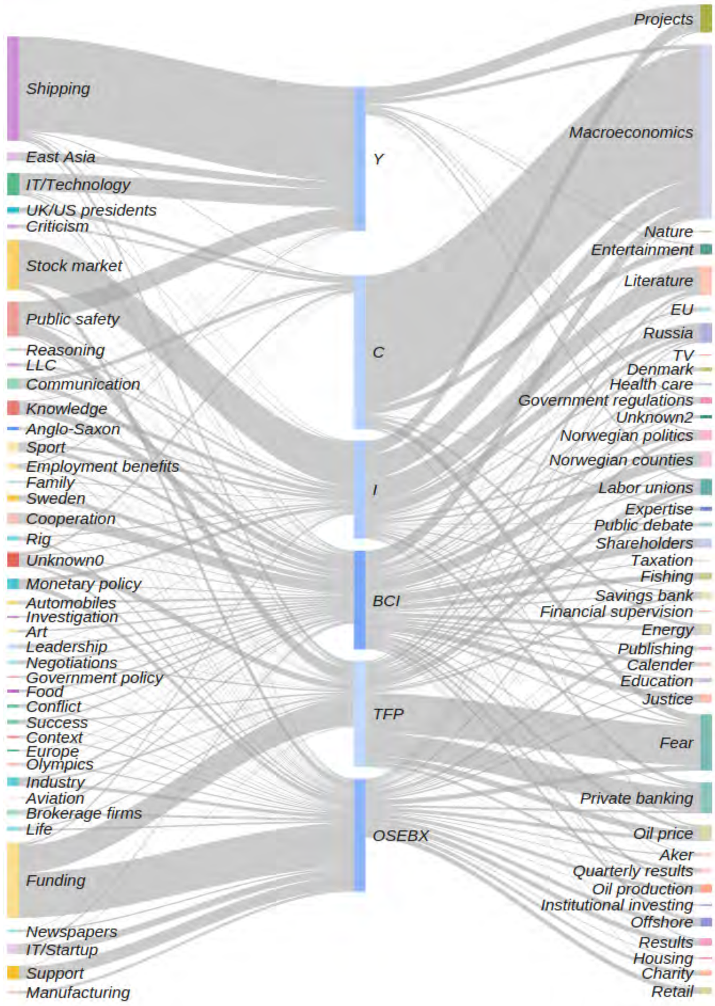


Figure 12. The Sankey diagram reports all topics and their associated weight in predicting a particular outcome variable. The thickness of the line connecting a topic and an outcome variable indicates the value of the weight. A thicker line means a higher weight. The weights are computed as described by equation (5) in Section 2.3. By construction, the results reported here are a direct function of those reported in Figure 3. This figure gives a clearer picture of which topics receive a weight and not, while Figure 3 gives a better picture of how predictable (using topics) the different outcome variables are relative to each other.

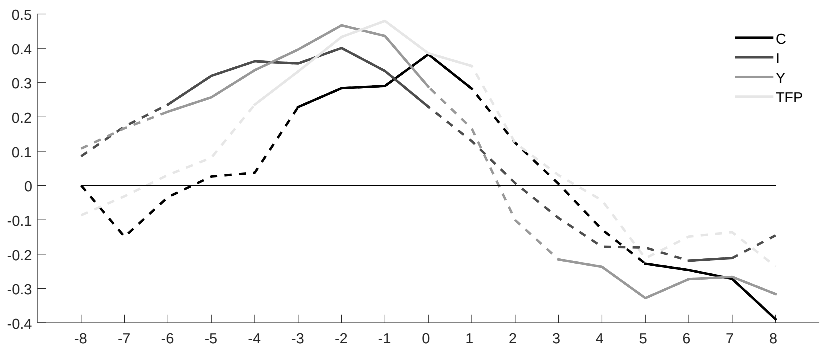


Figure 13. The graph reports the autocorrelation coefficient, at lead and lags, between the aggregated news index and four measures of economic activity: consumption (C), investment (I), output (Y) and total factor productivity (TFP), all transformed to year-on-year growth rates. A solid line indicates that the autocorrelation is significant ($\alpha = 0.05$).

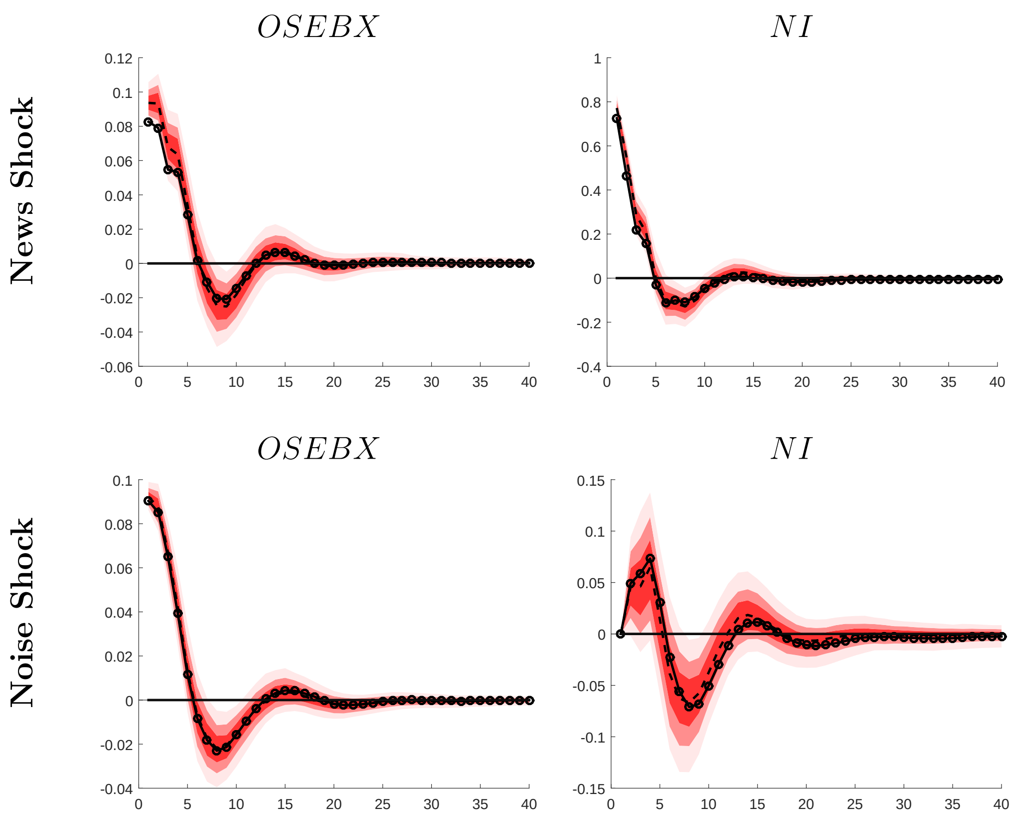


Figure 14. Impulse responses Benchmark model: Asset prices and the news index. See also Figure 5.

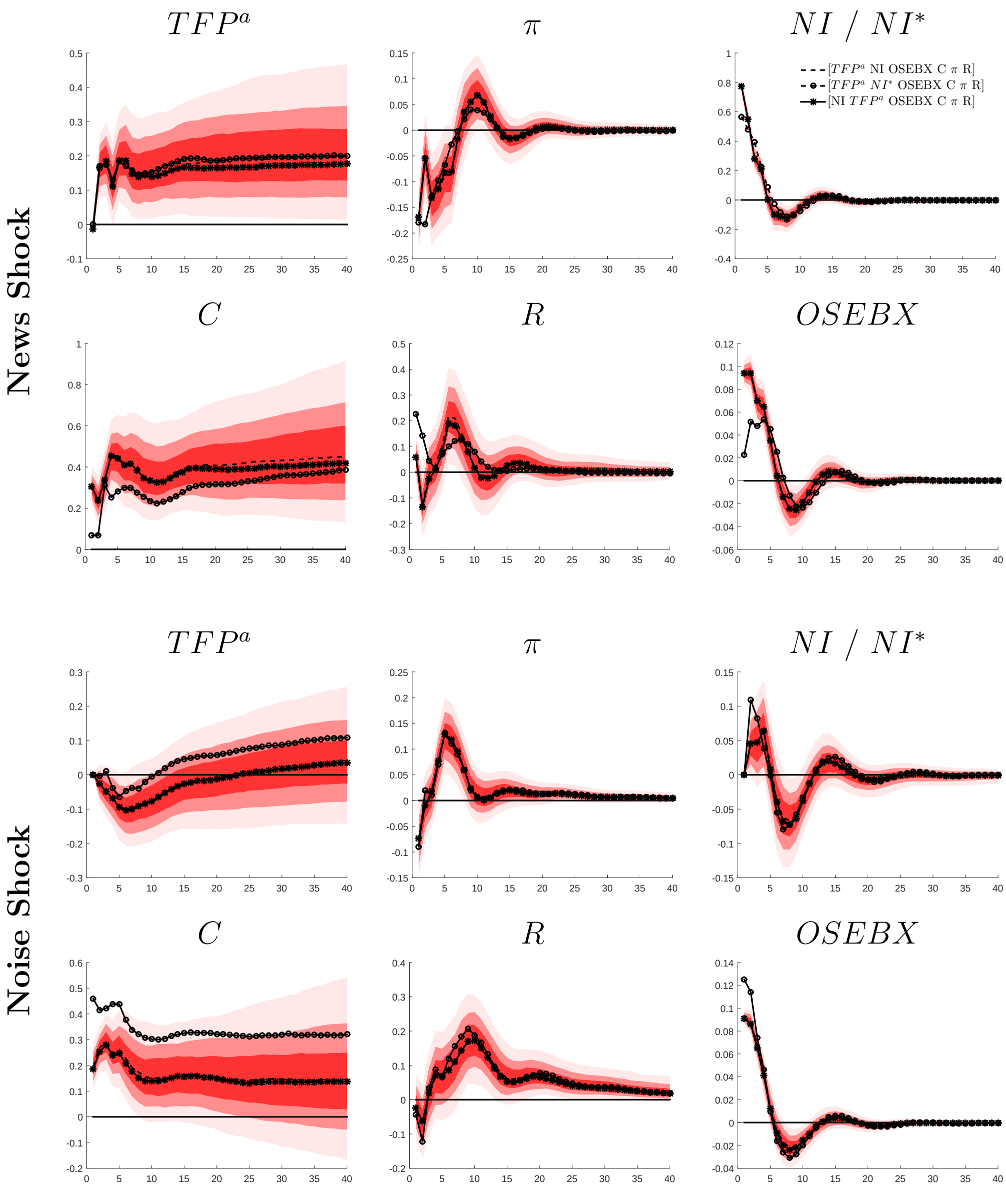


Figure 15. Robustness: Impulse responses Benchmark model, alternative news index, and alternative ordering. Each graph reports the response (in percent) to an initial one standard deviation shock across response horizons. The color shadings represent the 70, 50, and 30 percent quantiles of the posterior distribution for the Benchmark model. The black dotted line is the associated median estimate. See also Figure 5. The alternative models are described by the legend in the upper right corner. The NI^* variable is an aggregated news index constructed based on how well the news topics predict future TFP.

Table 4. LDA model evaluation: Number of topics and Perplexity score. See also the discussion in Appendix G.1.

	20	**40**	**60**	**80**
Perplexity score	6630	5484	4853	4409

Appendix G Models

In the analyses three types of models are employed. To extract the topics we use a Latent Dirichlet Allocation (LDA) model. To run the predictive regressions and construct the news indexes we use a Latent Threshold model (LTM). Finally, to explore the causes of economic variation we employ a Bayesian Vector Autoregression (BVAR). We describe these models in greater detail below.

G.1 Latent Dirichlet Allocation Model

The LDA model was developed in Blei et al. (2003). We implement the algorithm as described in Griffiths and Steyvers (2004). First, define T as the number of topics and:

$$P(w_i) = \sum_{j=1}^{T} P(w_i|z_i = j)P(z_i = j), \tag{13}$$

as the probability of word i occurring in a given document. Here, w_i is word i, and z_i is a latent variable denoting which topic word i is drawn from. The term $P(w_i|z_i = j)$ denotes the probability that word i is drawn from topic j, while the last term, $P(z_i = j)$, gives the probability of drawing a word from topic j in the current document. Different documents will have different probabilities for drawing words from the various topics.

Let D be the number of documents in the corpus and W the number of unique words. The importance of the words for the different topics can then be represented as:

$$P(w_i|z = j) = \phi_w^{(j)}, \quad \text{for all} \quad j \in [1, T] \quad \text{and} \quad w_i \in \{w_1, w_2, \ldots, w_W\} \tag{14}$$

where ϕ is a set of T multinomial distributions over the W words. The importance of a topic within a given document is defined as:

$$P(z = j) = \theta_j^{(d)}, \quad \text{for all} \quad j \in [1, T] \quad \text{and} \quad d_i \in \{d_1, d_2, \ldots, d_D\} \tag{15}$$

where θ is a set of D multinomial distributions over the T topics.

Given D, T, and W, the goal is to obtain estimates of ϕ and θ that maximizes equation (13) for all i, i.e., the probability that a word appears in the corpus. However, this approach is susceptible to problems involving local maxima and slow convergence. We follow Griffiths and Steyvers (2004) and instead use Bayesian estimation and Gibbs simulations. This strategy for discovering topics does not treat ϕ and θ as parameters to be estimated, but instead tries to approximate the posterior distribution over the assignments of words to topics, $P(z|w)$. Estimates of ϕ and θ are then obtained by examining the posterior distribution.

A complete probabilistic representation of the LDA model is:

$$w_i|z_i, \phi^{(z_i)} \sim \text{Discrete}(\phi^{(z_i)}) \tag{16a}$$

$$\phi \sim \text{Dirichlet}(\beta) \tag{16b}$$

$$z_i|\theta^{(d_i)} \sim \text{Discrete}(\theta^{(d_i)}) \tag{16c}$$

$$\theta \sim \text{Dirichlet}(\alpha) \tag{16d}$$

where α and β are hyper-parameters specifying the prior distribution for ϕ and θ. Since these priors are conjugate, we can integrate them out of the joint distribution $P(w, z) = P(w|z)P(z)$, using the representation in (16), and use the resulting distribution to approximate the conditional posterior:

$$P(z|w) = \frac{P(w, z)}{\sum_{j=1}^{T} P(w, z_j)} \tag{17}$$

We refer to Griffiths and Steyvers (2004) for specific details on how this is done using Gibbs simulations, and on how estimates of ϕ and θ can be obtained from the posterior.

Before estimation we pre-define three parameters: the number of topics, T, and the two hyper-parameters of the Dirichlet priors, α and β. The two latter are defined as a function of T and the number of unique words:

$$\alpha = \frac{50}{T}, \quad \text{and} \quad \beta = \frac{200}{W}.$$

which also is the same prior specification as used in Griffiths and Steyvers (2004). Here we note that $W = 250834$.

Choosing the value for T is essentially a model selection problem. To evaluate the performance of the topic model for different number of topics we use the perplexity score (equivalent to the predictive likelihood), defined as follows:

$$\text{Perplexity}(w) = \exp\left\{-\frac{\mathcal{L}(w)}{W}\right\}, \tag{18}$$

where

$$\mathcal{L}(w) = \log P(w|z) \tag{19}$$

As seen in Table 4, the lowest perplexity score, i.e., the best model fit, is obtained for $T = 80$, which we use in our preferred LDA specification.

G.2 Model LTM

In Section 2.1 of the main paper, we describe the Latent Threshold Model (LTM). Here we provide the estimation details. For convenience we first repeat the model, which can be written as follows:

$$y_t = x'_{t-1} b_t + u_t \qquad u_t \sim N(0, \sigma_u^2) \tag{20a}$$

$$b_t = \beta_t \varsigma_t \qquad \varsigma_t = I(|\beta_t| \geq d) \tag{20b}$$

$$\beta_t = \Xi \beta_{t-1} + e_t \qquad e_t \sim N(0, \Sigma_e) \tag{20c}$$

where t is the time index, x_{t-1} is a $(n \times 1)$ vector of (lagged) variables used for prediction, b_t a $(n \times 1)$ vector of time-varying parameters. ς_t is a zero one variable, who's value depends on the indicator function $I(|\beta_t| \geq d)$. If the ith element in $|\beta_t|$ is above the ith element in the $(n \times 1)$ threshold vector d, then $\varsigma_t = 1$, otherwise $\varsigma_t = 0$. e_t is a $(n \times 1)$ vector of disturbances associated with the time-varying parameters. We assume that e_t and u_t are independent. Apart from equation (20b), the system in (20) has a standard state space form.

To simulate from the conditional posterior of β_t and d in (20b), we follow the procedure outlined in Nakajima and West (2013). That is, conditional on all the data and hyper-parameters in the model, x_T, d, Σ_e and σ_u^2, we draw the conditional posterior of β_t sequentially for $t = 1 : T$ using a Metropolis-Hastings (MH) sampler. As described in Nakajima and West (2013), the MH proposals come from a non-thresholded version of the model specific to each time t, as follows. Fixing $\varsigma_t = 1$, takes proposal distribution $N(\beta_t | m_t, M_t)$ where:

$$M_t^{-1} = u_t^{-2} x_{t-1} x'_{t-1} + \Sigma_e^{-1}(I + \Xi'\Xi) \tag{21a}$$

$$m_t = M_t[u_t^{-2} x_{t-1} y_t + \Sigma_e^{-1}\{\Xi(\beta_{t-1} + \beta_{t+1}) + (I - 2\Xi + \Xi'\Xi)\beta_0\}] \tag{21b}$$

for $t = 2 : T-1$. For $t = 1$ and $t = T$, a slight modification is needed. Details can be found in Nakajima and West (2013). The candidate is accepted with probability:

$$\alpha(\beta_t, \beta_t^p) = min\left\{1, \frac{N(y_t|x'_{t-1}b_t^p, u_t^2)N(\beta_t|m_t, M_t)}{N(y_t|x'_{t-1}b_t, u_t^2)N(\beta_t^p|m_t, M_t)}\right\} \tag{22}$$

where $b_t = \beta_t\varsigma_t$ is the current state, and $b_t^p = \beta_t^p\varsigma_t^p$ is the candidate.

The independent latent thresholds in d can then be sampled conditional on the data and the hyper-parameters. For this, a direct MH algorithm is employed. Let $d_{-j} = d_{0:s}\backslash d_j$. A candidate is drawn from the current conditional prior, $d_j^p \sim U(0, |\beta_0| + K)$, where K is described below, and accepted with probability:

$$\alpha(d_j, d_j^p) = min\left\{1, \Pi_{t=1}^T \frac{N(y_t|x'_{t-1}b_t^p, u_t^2)}{N(y_t|x'_{t-1}b_t, u_t^2)}\right\} \tag{23}$$

where b_t is the state based on the current thresholds (d_j, d_{-j}), and b_t^p the candidate based on (d_j^p, d_{-j}).

Lastly, conditional on the data, the hyper parameters and the time-varying parameters, we can sample σ_u^2 and Σ_e using the standard inverse Gamma and Wishart distributions, respectively. For each of these elements we use a degrees of freedom prior of 10, and set the prior variance to 0.01.

The K parameter, used to draw d_j^p, controls our prior belief concerning the marginal sparsity probability. For example, assuming that a time-varying parameter follows $B_t \sim N(0, v^2)$, and marginalizing over B_t, it can be shown that $Pr(|B_t| = 0) = 2\Phi(\frac{d}{v}) - 1$, where Φ is the standard normal CDF. Defining $K = \frac{d}{v}$ as the standardized scaling parameter with respect to the threshold, it can be seen that $K = 3$ implies a marginal sparsity probability exceeding 0.99. As described in Nakajima and West (2013), a neutral prior will support a range of sparsity values in order to allow the data to inform on relevant values, and they suggest that setting $K = 3$ is a reasonable choice.[37]

In essence, the MH steps described above are identical to those proposed and described by Nakajima and West (2013). We only differ in the assumptions we make about the process for β_t. Here we assume that they follow Random Walks, thus, Ξ is an identity matrix, while Nakajima and West (2013) assume that they follow AR(1) processes. The Random Walk is non-stationary, and

[37]Note that when combined with the priors over the other hyper-parameters in the model, the implied marginal prior for each threshold will not be uniform, see Nakajima and West (2013) for details.

does not have a marginal distribution. For this reason the K parameter is also set differently in our application than theirs. We set $K = 1$, which, based on the reference model without latent threshold dynamics, seems to be a reasonable prior. Finally, the prior mean and covariance for the initial states are set to zero and the identity matrix, respectively.

G.3 Model BVAR

The VAR model was introduced in Section 3 of the main paper. We repeat it here for convenience:

$$y_t = \alpha + \phi_1 y_{t-1} + \ldots + \phi_p y_{t-p} + u_t \tag{24}$$

where y_t is a $(n \times 1)$ vector of endogenous variables, α is a $(n \times 1)$ vector of constants, ϕ_i, for $i = 1, \ldots, p$, is a $(n \times n)$ matrix of parameters. u_t are the reduced form residuals, with covariance $E(u_t u_t') = Q$.

For notational purposes it is helpful to put the VAR in SUR form. By abusing notation we define:

$$y = X\beta + \epsilon \tag{25}$$

where $y = [y_1, \cdots, y_T]'$, $X = [X_1, \cdots, X_T]'$, $\epsilon = [\epsilon_1, \cdots, \epsilon_T]'$ and $\beta = [\beta_1, \cdots, \beta_n]'$, with $\beta_k = [\phi_{1,k}, \cdots, \phi_{p,k}]$ for $k = 1, \ldots, n$. Further,

$$X_t = \begin{pmatrix} x_{t,1} & 0 & \cdots & 0 \\ 0 & x_{t,2} & \ddots & \vdots \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \cdots & \cdots & x_{t,q} \end{pmatrix}$$

with $x_{t,k} = [y_{t-1}', \cdots, y_{t-p}']$. Finally, $\epsilon \sim i.i.d.N(0, I_q \otimes Q)$.[38]

We simulate the SUR system sequentially using Gibbs simulations and the Normal-Wishart prior:

$$p(\beta, Q) = p(\beta)p(Q^{-1}) \tag{26}$$

where

$$p(\beta) = f_N(\beta | \underline{\beta}, \underline{V}_\beta) \tag{27}$$

$$p(Q^{-1}) = f_W(Q^{-1} | \underline{v}_Q, \underline{Q}^{-1}) \tag{28}$$

[38] With the VAR specified in SUR form it becomes easy to adjust the VAR(p) model such that different regressors can potentially enter the n equations of the VAR(p).

To avoid over-fitting, $\underline{\beta}$, $\underline{V}_\beta$, and $\underline{Q}^{-1}$ are set in a Minnesota style fashion, see, e.g., Koop and Korobilis (2010). That is, $\underline{\beta}$ for each dependent variable is set at its univariate AR estimate, and zero everywhere else. $\underline{V}_\beta$ is a diagonal matrix where each element is a scaled measure of the variance associated with the AR equation estimate. For lags of the dependent variable itself we use a scale of 1; for other lags we use a scale of 0.4. For exogenous measures, i.e., the constant, we use 0.3. $\underline{Q}^{-1}$ is set equal to its initial OLS estimate. Lastly, we set $v_Q = 30$, reflecting our relatively uninformative view on what the parameters of the VAR should be.

Based on these priors the conditional posterior of β is:

$$\beta|y, Q^{-1} \sim N(\overline{\beta}, \overline{V}_\beta)_{I[s(\beta)]} \tag{29}$$

with

$$\overline{V}_\beta = (\underline{V}_\beta^{-1} + \sum_{t=1}^{T} X_t' Q^{-1} X_t)^{-1} \tag{30}$$

and

$$\overline{\beta} = \overline{V}_\beta(\underline{V}_\beta^{-1}\underline{\beta} + \sum_{t=1}^{T} X_t' Q^{-1} y_t) \tag{31}$$

$I[s(\beta)]$ is an indicator function used to denote that the roots of β lie outside the unit circle.

The conditional posterior of Q^{-1} is:

$$Q^{-1}|y, \beta \sim W(\overline{v}_Q, \overline{Q}^{-1}) \tag{32}$$

with

$$\overline{v}_Q = \underline{v}_Q + T \tag{33}$$

and

$$\overline{Q} = \underline{Q} + \sum_{t=1}^{T}(y_t - X_t\beta)(y_t - X_t\beta)' \tag{34}$$

G.3.1 Identification

The mapping from the reduced form residuals, u_t in equation (24) to the structural innovations, e_t, is in this paper obtained using a Cholesky decomposition of $E(u_t u_t') = Q$, such that $u_t = A_0 e_t$. From this it follows that:

$$\begin{bmatrix} u_{1,t} \\ u_{2,t} \\ \vdots \\ u_{n,t} \end{bmatrix} = \begin{bmatrix} a_{11} & 0 & 0 & 0 \\ a_{21} & a_{22} & 0 & 0 \\ \vdots & \vdots & \ddots & 0 \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{bmatrix} \begin{bmatrix} e_{1,t} \\ e_{2,t} \\ \vdots \\ e_{n,t} \end{bmatrix} \tag{35}$$

where e_t are the structural disturbances, with $\Sigma_e = I$, such that $Q = A_0 A_0'$.