University of Tartu

Faculty of Social Sciences

School of Economics and Business Administration

Caio Augusto Raduy Koslyk

NEWS-DRIVEN BUSINESS CYCLES: A NARRATIVE APPROACH

Master's thesis

Supervisor: Lenno Uusküla (PhD)

Tartu 2019

Name and signature of supervisor...

Allowed for defense on..

(date)

I have written this master's thesis independently. All viewpoints of other authors, literary sources and data from elsewhere used for writing this paper have been referenced.

. . . .

(signature of author)

News-Driven Business Cycles: A Narrative Approach

Caio Augusto Raduy Koslyk[1]

Abstract

This paper analyses the effects of technology news on the US business cycle. The paper suggests a new frequency-based index about the technology news from a major news outlet for the period 1948Q1 to 2017Q3. The obtained index is then used within a structural vector autoregressive framework. An increase in technology news has no effect on productivity in the short run, but is correlated with strong productivity growth in the longer run. The news on technology leads to an immediate and persistent increase in consumption, investments and hours worked and explains a high share of the forecast error variance decomposition at business cycle frequency.

JEL codes: C8, D84, E32, O33

Keywords: technology news, textual analysis, VAR

[1]E-mail: caiokoslyk@gmail.com

Acknowledgments

I would like first to express my deepest gratitude to my supervisor Lenno Uusküla for the valuable comments, discussions and guidance, which were certainly fundamental for the development of this thesis. I would like to extend my sincere thanks to Diana Gabrielyan for her meticulous review of the thesis and insightful suggestions on how to improve it. Still, any errors are my own and should not tarnish the reputations of these esteemed persons.

Furthermore, I must also thank Jaan Masso and Kadi Kaur for their help and patience during the whole period of studying. I gratefully acknowledge the unfaltering moral support of my friends, both old and new, throughout my educational journey at the University of Tartu.

Finally, I am extremely grateful to my family, especially my parents and my brother, albeit distant, for their unconditional love and support along the way.

Contents

1 Introduction

Can news about future productivity drive economic fluctuations? Since early 20th century, economists have been trying to study the role of changes in agents' expectations on business cycles. Keynes (1936) discussed how spontaneous optimism can cause economic cycles, as animal spirits, rather than mathematical expectations are the main cause for our action. Already Pigou (1927) had stressed the importance that errors of optimism and pessimism of businessmen can have in creating trade cycles. Expectations still play an important role in modern macroeconomics and are central components of the most prominent schools of thought.

More recently, reignited by the influential paper by Beaudry and Portier (2006), many economists have assessed whether expectations can cause business cycles: do changes in information today due to news about future fundamentals cause comovements between the main macroeconomic variables? Most of the studies focus on the identification of these news shocks, anticipated shocks to fundamentals in the future, through innovations to variables that might reflect agents' expectations and signals about future developments, such as stock prices and confidence indexes, see for example Barsky and Sims (2012). Beaudry and Portier (2014) and Ramey (2016) present a comprehensive summary of the news shock literature.

This paper studies a direct and widely spread source of information about technological developments which are newspaper articles on technology. Using an extensive news corpus from a major US newspaper – The New York Times – the paper calculates a value of technology news index (TNI) for each quarter. It captures the intensity in which technology articles are being written, as they might contain information about future expansions in the productivity frontier.

The paper then uses the newly constructed index in a VAR model with total factor productivity (TFP) of Fernald (2014) and macroeconomic aggregates. The model is estimated for the period 1948Q1 to 2017Q3. The identification scheme for the benchmark model follows Beaudry and Portier (2006) strategy of combining short and long-run restrictions. Specifically, the news shock is identified as an innovation to the TNI that does not affect TFP on impact but has a permanent effect on the long-run, while innovations to TFP are unrestricted both on the short and long-run.

The paper finds that the constructed TNI using newspaper articles carries information about future productivity given that it is able to produce a S-shaped response of TFP to a news shocks. Initial effect on the TFP is very small or even negative and the positive TFP developments appear three to ten years later. The structural shocks to the index produces comovements between macroeconomic aggregated that precedes growth in productivity. Consumption, investments, hours and output react strongly and positively right after the shock. The shocks are able to explain substantial share of fluctuations at business cycles frequencies, accounting for more than half of the cyclical forecast variance of macroeconomic data.

The results are also robust to a different measure of productivity, namely labour productivity, producing similar patterns to the benchmark model. Additionally, the shock identified using the TNI, is almost perfectly collinear and produces similar dynamics as an innovation to TFP that drives its long run path, consistent with the results in Beaudry and Portier (2006) using stock prices.

The paper contributes to the growing literature of textual analysis, following Grimmer and Stewart (2013), Shiller (2017) and Gentzkow et al. (2017). The paper suggests an alternative approach to identify technology news shocks, using narrative analysis and textual data. It is closely related to the traditional news shocks literature, as such shocks are regarded as anticipated shocks to macroeconomic fundamentals of some future period. Differently the identification of news shocks here relies on an index constructed from a primary source of technology news.

Larsen and Thorsrud (2019b) hypothesizes that the more intensive a given topic is discussed in the newspaper, the more likely is that this topic represents something of importance for the economy's future. In light of this, the construction of our intensity-based index is simply the relative importance that articles associated with a list of technology terms and keywords are published in The New York Times in a given period.

Alternative measures of news, such as innovations to asset prices and confidence index might not be the best variables to identify technology news shocks, because 1) they might fluctuate because of fundamental information other than technological advances; 2) they can fluctuate due to nonfundamental information, reflecting sentiments or animal spirits 3) it was shown that they tend to over or under-react to news publications, see Larsen and Thorsrud (2019b) for an extensive discussion of the topic.

Second, it contributes to the literature by providing additional evidence in favour of a news-driven business cycles on the US economy. Positive comovements are in line with models with variable capital utilisation, adjustment costs to investment, small short-run wealth effects on the labour supply (Jaimovich and Rebelo, 2009). Under these assumptions, the substitution effect is greater than the income effect in the labour market, hence, due to positive news shocks, agents expect that their income will be higher in the future, therefore, they consume more today, but, as leisure becomes more expensive, they supply more labour. Empirical papers that support such situation are Beaudry and Portier (2006), Beaudry and Lucke (2010), Schmitt-Grohé and Uribe (2012), and Larsen and Thorsrud (2019b). Negative comovements between consumption and hours worked are in line with the traditional neoclassical models and the permanent income hypothesis, so that due to a greater wealth effect over the substitution effect, agents can afford more leisure today due to expected higher income, increasing agents' consumption and inducing a decline in hours worked, empirical paper supporting this idea are Barsky and Sims (2011), Barsky et al. (2014) and Forni et al. (2014).

The rest of this paper is organized as follows. Section 2 gives a review of the literature on news shocks. Section 3 proposes a simple theoretical news-driven business cycles model and presents the identification strategy used later in the paper. In Section 4 discusses the construction of technology news and other macroeconomic data. Section 5 presents results of model estimations along with discussions regarding the relation of our narrative-based news shocks to standard literature on news shocks. Finally it also presents extensive robustness checks of results that are performed. Section 6 concludes the paper.

2 Literature review

The view that changes in expectations about future fundamentals are potential sources of economic fluctuations has been discussed for long. But Barro and King (1984) claimed that unless expectational changes are not accompanied by actual changes in current productivity, these type of shocks are not able to generate positive comovements of consumption, worked hours, and output.

The debate about news-driven business cycles became popular again after the publication of a series of seminal articles by Beaudry and Portier (2004, 2005, 2006). These papers provided formal foundations for the role of expectations on macroeconomics fluctuations and evidences that a shock that does not affect productivity in the short run but affects it in the long run – technology news shock – can be the main driver of business cycles.

A rich literature on news shocks followed the publication of those papers, and in this section, we present a review on the main works on this area of research. For that, the literature review is divided into theoretical and empirical developments.

Business cycles data shows that there are both an aggregate and sectoral comovements of variables such as output, consumption and investment. Cochrane (1994) claims that variants of the neoclassical models are unable to generate booms in response to expectation of higher future TFP. Due to the permanent income hypothesis, good news about future productivity creates a wealth effect on agents, they increase their consumption, and substitute work for leisure, reducing the labour supply, in consequence, output fall as well. Hence, good news about tomorrow would generate a recession already today (Jaimovich and Rebelo, 2009, p. 1099).

In light of the weaknesses of traditional real business cycle (RBC) models to generate macroeconomic aggregates fluctuations consistent with the business cycles data, Beaudry and Portier (2004) proposes a formalization of a Pigouvian view of business cycles in which busts and booms may be caused by forecast errors of agents regarding the future conditions of the economy. Economic agents make decisions based on newly arrived and noisy information: when they are optimistic about the future condition of the economy, they decide to make investments in capital to meet future demand, hence, consumption, hours worked, investment, and price of capital positively comove.

When future arrives and there are material technological improvements, the economy keeps growing. However, when expectations of agents are not fully met, they realize that there is an overinvestment in the economy, in consequence, there is a significant reduction in employment followed by a period of retrenched investment with potential recessionary effects. All in all, boom and recessions may arise due to expectation and realization of technological growth even if technology does not actually change.

Jaimovich and Rebelo (2009) propose a model that addresses the issues with RBC models by adding several features to the traditional neoclassical model: variable capital utilization, adjustment costs to investment and a weak short-run wealth effect on the labour supply. As a result, the model generates comovements in macroeconomic aggregates following TFP contemporaneous and news shocks, and investment-specific technical change.

When it comes to empirical results, the seminal paper on the modern news shocks literature is the empirical analysis of US business cycles by Beaudry and Portier (2006). They study how stock price movements, as seen as changes in expectations about the future economic conditions, in conjunction with movements in TFP, favour the view that a shock reflecting news about future technological opportunities that does not have effect on productivity in the short run, but do effects it with substantial delay can be a major source of economic fluctuations.

For that purpose, by assuming that only technology and demand shocks drive business cycles, they analysed a bivariate SVAR system formed by TFP and stock prices, and sequentially imposed two different restrictions schemes in order to identify the shocks: in the first one, the demand shock is identified through a short-run restriction where an innovation to stock prices that is orthogonal to TFP has no contemporaneous impact on TFP; and in the second one, the technology shock is identified by imposing long run restrictions. Surprisingly, they found out that both shocks, which by construction were supposed to be demand and technology shocks, respectively, led to similar results in the system. Hence, given their identification strategies, both shocks could be considered news shocks.

By extending for higher dimension systems and using a combination of short and long-run restrictions, they found out that this shock, seen as news about future technological conditions, generates comovements in macroeconomic aggregates and plays a major in the US economy. It is responsible for about 50% of business cycles fluctuations.

Beaudry and Lucke (2010) find support for the Beaudry and Portier (2006) evidence using similar identification schemes with combinations of short and long-run restrictions. Instead their system is augmented allowing for five different shocks, and find that news shocks are the most important driver of macroeconomic fluctuations. See Beaudry and

Portier (2014) for a review on the literature of news shocks and updates of the Beaudry and Portier (2006) data for the sample 1947Q1–2012Q1.

However, the identification scheme used in Beaudry and Portier (2006) had been target of criticism by some authors. Kurmann and Mertens (2014) claim that the approach by Beaudry and Portier is only valid for their bivariate system, whereas for the higher-dimensional systems, the VECM with long run restrictions underidentifies the system, yielding a large set of possible solutions, hence no unique solution can be found.

Barsky and Sims (2011) proposes a novel identification scheme in VAR: news shocks are identified as the linear combination of disturbances that are orthogonal to the TFP innovation which maximizes the sum of contributions to the forecast error variance of TFP over a ten year horizon. Their results show that a news shock is disinflationary on impact, and increases consumer confidence and stock prices immediately. Moreover, at the time news shocks arise, consumption increases and keep rising further over time, however, output, hours and investment jump down on impact before recovering after a few quarters when TFP starts to increase. Such findings are consistent with the predictions of traditional neoclassical models, hence, there are no significant output booms following news shocks, a contrasting results with those in Beaudry and Portier (2006).

Barsky et al. (2014) found similar results using a similar identification scheme, but maximizing TFP's FEV at a 20-quarter horison, using a 9-variable VAR system. Also under a maximum variance scheme, Forni et al. (2014) estimates a factor-augmented structural VAR, and find that, although news shocks play a non negligible role in output fluctuations, consumption and output increase considerably, while hours worked fall at the moment when news arrive, in line with traditional neoclassical models.

Schmitt-Grohé and Uribe (2012) estimates a dynamic stochastic general equilibrium (DSGE) framework augmented by the features present in Jaimovich and Rebelo (2009). Allowing for seven conventional shocks, and correspondent news shocks, the results are in line with Beaudry and Portier (2006), as they find that news shocks account for about half of the variation in output, consumption, investment and employment.

From the new instrumental variable (IV) SVAR literature (Stock and Watson, 2018) arises two papers with applications to news shocks: Miranda-Agrippino et al. (2018), using US utility patent applications to construct an external instrument for the identification of such shocks, and Cascaldi-Garcia (2019), using an instrumental variable based on forecast revisions from the Survey of Professional Forecasters (SPF) to identify news shocks. The findings in both paper are very similar and partially support the findings in Beaudry and Portier (2006): news shocks can induce macroeconomic comovements, however, their role is rather limited, as they are non-trivial drivers of fluctuations at business cycles frequencies.

The closest paper to the approach in this paper is the pioneering work of Larsen and Thorsrud (2019b) that provides a narrative-based approach to news shocks through news topics extraction from a major Norwegian business newspaper using textual analysis and natural language processing (NLP) methods. Larsen and Thorsrud (2019b) construct an aggregate news index containing the news topics with the highest predictive power of stock prices for each period t, the so-called financial news index (FNI), to be used on the identification of technology news shocks in a structural VAR analysis with Choleski decomposition. Their results showed that, following a news shock, both consumption and employment increase on impact and remain positive for longer horizons, inflation falls and, as expected, asset prices increase. Hence, news shocks are important sources of business cycles fluctuations in the Norwegian economy.

3 Method

3.1 A simple model of news-driven business cycles

This subsection presents the simple model to show how anticipated shocks to economic fundamentals can drive business cycles as in Forni et al. (2017). It helps to estimate an empirical model and put estimated results in the context of economic theory.

We assume that economic fundamentals are composed by the their own past values and driven by a component with features of a *news* shock, a shock with a delayed and permanent effects on that fundamental. Agents receive information about this shock at the time when it arises. However, the information received by the agents is noisy, it is composed by the news shock, and by a temporary shock, which does not affect fundamentals, also known as the *noise* shock.

Consequently, agents cannot recognize the true shock when they receive the signal, and face a signal extraction problem. Nonetheless, they form their expectations about future economic fundamentals using their current information set, a noisy signal about long-run fundamentals and the realized levels of these fundamentals. Based on these expectations, agents make their current consumption decisions, creating macroeconomic fluctuations before any actual changes in fundamentals.

Formally, we assume that the economic fundamental of interest is productivity, a_t, and that it follows an exogenous dynamic process composed by its own past level and a structural shock with delayed effect, the so-called technology news shock:

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

where ϵ_{t-1} is serially uncorrelated with a_{t-1}, and represents the *news* shock, a shock that occurs at $t-1$ and only affects a with one-lag delay, at time t, and follows an i.i.d. Gaussian process with mean zero and variance σ_ϵ^2.

However, agents cannot observe the pure *news* about future productivity when it arises, because the signal they receive is noisy and also contains a *noise* term, thus they are not able distinguish between noise and true shocks at the period when they receive the information.

In the case of *news* shock ϵ_{t-1}, it is only recognizable one period later, at time t, because agents observe the change in productivity at t, which is exclusively driven by *news* from the previous period, as shown in Equation 1.

Generalising, the noisy signal, s_t, received by the agents at any period t is given by the Equation 2. It is composed by the *news* shock ϵ_t that affects a in the next period, and a noise component that has no effects on productivity, v_t, which is a Gaussian white noise with variance σ_v^2, uncorrelated with ϵ_t. Furthermore, the variance of s_t is simply the sum of the variances of ϵ_t and v_t, i.e., $\sigma_s^2 = \sigma_\epsilon^2 + \sigma_v^2$.

$$s_t = \epsilon_t + v_t. \tag{2}$$

Agents recognize this signal, s_t, at the time when it arises, and, as mentioned before, the structural component of the signal is only revealed one period later, when the agents observe the gains in productivity due to the delayed effects of the *news* shock, as seen in Equation 1. When agents gather these two observations, they become able to determine the magnitude of the *noise* shock as well through a simple subtraction exercise, $v_t = s_t - \epsilon_t$, following Equation 2.

Hence, at time t, the information set of the agents, I_t, is composed by the realized values of a, past and current values of the noisy signal s_t, and past values of *news* and *noise*. This information set is then used by agents to form their expectations about the future and decide on consumption today.

For consumption c_t we assume that it is set at t based on agents' expectations about future productivity, given their current information set, as shown in Equation 3.

$$c_t = \lim_{j \to \infty} E(a_{t+j} | I_t), \tag{3}$$

Moreover, we assume that output, y_t, is fully determined by demand, hence $y_t = c_t$. Employment adjusts to clear the labour market, so that labour input, l_t, is equal to the subtraction of a_t from y_t.

The dynamic process of a_t, given by Equation 1, and the expected value of ϵ_t implies that $E(a_{t+j}|I_t) = E(a_{t+1}|I_t)$ for any given $j > 1$. Therefore, we can rewrite Equation 3 as:

$$c_t = E(a_{t+1}|I_t). \tag{4}$$

From Equation 1, we take expectations, and find that the expected value of a_{t+1} is the expected value of the productivity at t, a_t, plus the expected value of ϵ_t given their current information set at period t, I_t. The $E(a_t|I_t)$ is equal to the current productivity,

a_t. Hence, we can rewrite Equation 4 as:

$$c_t = a_t + E(\epsilon_t | I_t). \tag{5}$$

However, we know that lagged values of a_t and s_t alone are not able to reveal ϵ_t, then we assume that $E(\epsilon_t|I_t)$ is simply the projection of ϵ_t on s_t, thus a function of the known values of the variances of *news* and *noise*:

$$E(\epsilon_t | I_t) = \gamma s_t, \tag{6}$$

where γ is equal to the signal-to-noise ratio, σ_e^2/σ_s^2.

Hence, from Equations 5 and 6, we have that c_t is determined by the current level of productivity and the noise signal that agents receive at t:

$$c_t = a_t + \gamma s_t \tag{7}$$

Substituting γs_t in Equation 7 for $\epsilon_t + v_t$, from Equation 2, results that the level of consumption at t can be rewritten as a function of current *news* and *noise*, as seen in Equation 8.

$$c_t = a_t + \gamma(\epsilon_t + v_t). \tag{8}$$

Therefore, if we take Equation 8 and apply the first difference operator in order to find the change in consumption, Δc_t i.e., the level of consumption today, c_t minus the level of consumption one period before, c_{t-1}, we have that:

$$\Delta c_t = \Delta a_t + \gamma \Delta(\epsilon_t + v_t) \tag{9}$$

From the Equation 1, we have that gains in productivity, Δa_t is equal to the news shock ϵ_{t-1}. Substituting this result in Equation 9, Δc_t becomes:

$$\Delta c_t = \epsilon_{t-1} + \gamma \Delta(\epsilon_t + v_t) \tag{10}$$

Applying the first different to both *news*, ϵ_t, and *noise*, v_t, we have that Δc_t becomes a function of past and current shocks, as shown in Equation 11.

$$\Delta c_t = \epsilon_{t-1} + \gamma((\epsilon_t - \epsilon_{t-1}) + (v_t - v_{t-1})) \tag{11}$$

Then, by redistributing the terms in Equation 11, we have a decomposition of changes in consumption fully-determined by news and noise shocks, and their variances, as seen in Equation 12.

$$\Delta c_t = \gamma\epsilon_t + (1-\gamma)\epsilon_{t-1} + \gamma(v_t) - \gamma(v_{t-1}). \tag{12}$$

Equivalently, for one-period-ahead, the changes in consumption Δc_{t+1}, which is equal to $c_{t+1} - c_t$, is determined as follows:

$$\Delta c_{t+1} = \gamma\epsilon_{t+1} + (1-\gamma)\epsilon_t + \gamma(v_{t+1}) - \gamma(v_t). \tag{13}$$

From Equation 12, it is possible to see that following a news shock ϵ_t that affects fundamentals only on the next period, and holding all other variables constant, c increases by $\gamma\epsilon_t$ on impact, jumping to a new level $c_t = c_{t-1} + \gamma\epsilon_t$.

At period $t+1$, following the news shocks at t, from Equation 13, consumption c increases by $(1-\gamma)\epsilon_t$, hence the change in consumption is $\Delta c_{t+1} = (1-\gamma)\epsilon_t$, and the new level of consumption is $c_{t+1} = (1-\gamma)\epsilon_t + c_t$, and substituting c_t, c_{t+1} becomes equal to $(1-\gamma)\epsilon_t + c_{t-1} + \gamma\epsilon_t$. Therefore, the accumulated effects of ϵ_t over the two periods, $\gamma\epsilon_t + (1-\gamma)\epsilon_t$, being equal to the magnitude of the news shock ϵ_t, c reaches its new long-run level, $c_{t+1} = c_{t-1} + \epsilon_t$.

Similarly, it is possible to do the exercise for a noise shock, v_t. From Equation 12, following a noise shock v_t, c increases by $\gamma(v_t)$ on impact, at time t, so the new level of consumption is $c_{t-1} + \gamma(v_t)$. For the following period, $t+1$, Equation 13 shows that v_t affects consumption negatively by $\gamma(v_t)$, hence c returns to its initial level c_{t-1}, because the total effect of v_t is $\gamma(v_t) - \gamma(v_t) = 0$ over the two periods.

It is worth noting that when both shocks, ϵ_t and v_t occur, agents cannot distinguish them. However, at $t+1$, because of the new level of a_{t+1}, agents can recognize that v_t was just noise, and undo the initial increase in consumption, reducing it by $\gamma(v_t)$.

To sum up, according to the model, both news and noise shocks can be sources of business cycles fluctuations without actual gains in productivity, but only the former has a permanent effect, while the latter has only a temporary effect on consumption and output.

We can also consider the situation in which the agents receive a perfect signal about future productivity. In this case v_t is always zero for any given t, and $\sigma_s^2 = \sigma_\epsilon^2$, so that the signal-to-noise ratio, σ_e^2/σ_s^2 is equal to unity. From Equation 12 and 13 one can see that consumption immediately jumps by ϵ_t and reaches its new long-run level $c_{t-1} + \epsilon_t$ right on impact.

So, when agents face imperfect information, i.e., $\gamma < 1$, they are more cautious to consume. And the higher the variance of noise, σ_v^2, the lower the γ, hence, the lower is the contemporaneous change in consumption after a news shock.

3.2 Identification

For the estimation of the effects of news shocks on the business cycle, the constructed technology news index is used to identify news shocks in a SVAR framework. The approach will closely follow Beaudry and Portier (2006) in a sense that technology news shocks are defined as those that have a delayed and permanent effect on productivity in the present case and where the news shocks is identified as an innovation to a textual-based index orthogonal to TFP as in Larsen and Thorsrud (2019b). Traditional macroeconomic variables are included in the system to estimate the effects of the news shock on the business cycles, allowing for a surprise technology shock, news shocks and a temporary shock to TFP.

Such identification scheme is justified by the following model of TFP for a trivariate system, with TFP and TNI placed first, present in Beaudry and Portier (2006).

$$TFP_t = R_t + D_t + \nu_t \tag{14}$$

$$R_t = R_{t-1} + \eta_{1,t} \tag{15}$$

$$D_t = \sum_{i=0}^{\infty} d_i \eta_{2,t-1} \;\; d_0 = 0, \;\; d_i \leq d_{t+1}, \lim_{i \to \infty} d_i = 1 \tag{16}$$

$$\nu_t = \rho \nu_{t-1} + \eta_{3,t}, \;\; where \;\; 0 \leq \rho < 1. \tag{17}$$

In words, the system given from Eq. 14 to 17 shows that TFP is driven by three components: a random walk, R_t, a diffusion process, D_t, and a temporary shock (or measurement error), ν_t. Therefore, the structural shock η_1 affects TFP on impact and has a permanent effect on it, and can be seen as a traditional technology shock, equivalent to place no restrictions on the $1,1$ elements of both impact and long-run matrices.

In Equation 16 the innovation η_2 is orthogonal to η_1 and has no contemporaneous impact on TFP. It takes some time to permanently affect productivity ($d_0 = 0$), hence, it illustrates a news shock, that it is expected to create an S-shaped response, in accordance with a technology diffusion process, so the element $1,2$ of the impact matrix is restricted to be zero, with no restrictions on the $1,2$ element of the long-run matrix; and, finally, η_3 is expected to have contemporaneous effect but a zero long-run effect on TFP, meaning that the element $1,3$ of the short-run matrix is unrestricted and the $1,3$ element of the long-run matrix is restricted to be zero.

The structural form of the VAR that describes the dynamics of a n x 1 vector y_t of endogenous variables is given by:

$$By_t = \alpha_0 + \sum_{i=1}^{p} \alpha_i y_{t-i} + \epsilon_t, \tag{18}$$

where α_0 represents all the deterministic parameters, such as constants and seasonal components, α_i are the autoregressive matrices of coefficients, B is the contemporaneous impact (structural) matrix, ϵ_t is an $n \times 1$ vector of independent structural innovations and p is the number of lags in the model. Furthermore, we assume that ϵ_t is normally distributed with mean zero and the covariance matrix is an identity matrix I, i.e., $E[\epsilon_t]$ $= 0$ and $E[\epsilon_t \epsilon_t'] = I$. Additionally, we also assumed that each shock is serially uncorrelated and independent of each other, $E[\epsilon_t \epsilon_s'] = 0$, for $s \neq t$.

By assuming the nonsingularity of B, the SVAR in Equation 18 can be rewritten in his reduced form as follows:

$$y_t = \delta_0 + \sum_{i=1}^{p} \delta_i y_{t-i} + A\epsilon_t, \tag{19}$$

where $A = B^{-1}$ and $\delta_j = B^{-1}\alpha_j$ for each $1 \leq j \leq p$. The reduced form innovations is related to the structural shocks in the following way $u_t = A\epsilon_t$, which implies that u_t also follows a normal distribution with mean zero, and covariance matrix $\Sigma_u = E[u_t u_t'] =$ AA'.

The reduced form VAR (Equation 19) can be rewritten as follows:

$$A(L)y_t = \delta_0 + u_t, \tag{20}$$

where $A(L) = (I_2 - \sum_{i=1}^{p} \delta_i L^i)$ and L represents the lag operator.

The moving average representation of y_t in Equation 20 can be achieved by premultiplying both sides of the equation by $A(L)^{-1}$ so that:

$$y_t = \mu_0 + \psi(L)u_t, \tag{21}$$

where $\psi(L) = (I_2 - \sum_{i=1}^{p} \delta_i L^i)^{-1} = \sum_{k=0}^{\infty} \psi_k L^l$, $\psi_0 = I_2$, $\psi_k = \delta_1^k$ and $\mu_0 = A(L)^{-1}\delta_0$.

By replacing u_t with $B^{-1}\epsilon_t$ in Equation 21, the structural moving average representation is recovered:

$$y_t = \mu_0 + \Theta(L)\epsilon_t, \tag{22}$$

where $\Theta(L) = \psi(L)B^{-1} = \sum_{k=1}^{\infty} L^k = B^{-1} + \psi_1 B^{-1} L^1 + ...$ and $\Theta_k = \psi_k B^{-1} = \delta_1^k$, for k $= 0,1,...$

As the elements of the matrices Θ_k give the dynamic multipliers of the vector y_t to structural shocks in ϵ_t, the long-run cumulative impact of the these shocks is captured by the matrix $\Theta(l) = \sum_{k=0}^{\infty} \Theta_k$, henceforth, called the long-run matrix.

For the structural interpretation of the innovations, it is necessary to recover the structural shocks from the reduced form innovations. However, given the assumed normality on ϵ_t, and the symmetry of the Σ_u matrix, to achieve a just identification of the system, it is necessary to impose $m(m-1)/2$ restrictions on the structural impact matrix B and on the long-run matrix $\Theta(l)$, in total.

In this paper, the identification scheme used follows the approach in Beaudry and Portier (2006) by imposing zero short and long-run restrictions in some of the components of B and $\Theta(l)$. Just-identification of news shocks can be achieved as follows for the trivariate where TFP_t and TNI are placed first in the system:

$$B = \begin{pmatrix} \star & 0 & \star \\ \star & \star & \star \\ \star & \star & \star \end{pmatrix} \qquad \Theta(l) = \begin{pmatrix} \star & \star & 0 \\ \star & \star & \star \\ \star & \star & 0 \end{pmatrix}.$$

And for the 4-variable specification:

$$B = \begin{pmatrix} \star & 0 & \star & 0 \\ \star & \star & \star & 0 \\ \star & \star & \star & 0 \\ \star & \star & \star & \star \end{pmatrix} \qquad \Theta(l) = \begin{pmatrix} \star & \star & 0 & \star \\ \star & \star & \star & \star \\ \star & \star & 0 & \star \\ \star & \star & \star & \star \end{pmatrix},$$

where $\star$ represents unrestricted parameters.

Under this identification scheme, given the order of the variables in the system mentioned above, the news shock, ϵ_2, is identified as the shock that does not affect productivity on the impact but it is able to affect contemporaneously all other variables, while it can affect TFP in the long-run along with the innovation in TFP, ϵ_1, and with a variable-specific shock, ϵ_4.

4 Data

4.1 Technology news index

This subsection presents a detailed description and rationale on the construction of the technology news index, and the next subsection describes other standard macroeconomic variables.

Not only is known that media coverage reflects the current and future state of the economy, but the relevant literature also suggests that it can affect public economic perceptions, as shown by Soroka et al. (2015), Damstra et al. (2018), and references therein. According to them, media coverage is seen as a filter for the voluminous economic data, in which the average citizen relies on to gauge changing economic conditions for expectations formation, hence it makes the newspaper-based approach a good strategy for assessing the signals about current developments that agents receive.

For the identification of technology news shocks, we construct an aggregate news index based on the idea that the relative frequency in which technology-related articles are published can give hints about - *current technological conditions*, and given the slow diffusion of innovations - the future condition of productivity.

For this purpose we use The New York Times news corpus. NYT is the second biggest newspaper by circulation in the US, covering wide array of topics. The data is retrived from the archive of its "The New York Times Developer Network"[2] that gives access to articles that dates back to 1851. The selection of the sample period is standard post WWII sample starting first quarter of 1948 and finishing in the third quarter of 2017. It is a relatively stable period in the US economy, has wide array of statistics available, and is used in other studies.

The selection of the articles for the analysis is an important issue as the population of articles is enormous and resources (time, computational) are scarce, so it is very important to rely on a selection criteria that might reflect actual news about science and technology. NYT classifies each article by assigning to it keywords that reflect its subject (henceforth, subject keywords).

In total, there are approximately 55,500 unique subject keywords in the dataset, and the process of filtering the sample of interest is made in three steps: 1) we need to reduce this universe of subject keywords, keeping only those related somehow with scientific and technological events, so a list containing "atemporal" technological and scientific terms and roots (as scientific and technological terminologies might frequently change over time) - precisely: *"scien", "tech", "new developments", "research", "new models", "inventions"* and *"patents"* was created; 2) a list of keywords containing at least one of the "atemporal" technological and scientific terms and roots from 1) was constructed (the full list of keywords is presented on the Appendix A); and 3) the articles associated with one or more subject keywords from the list are included in the sample. Table 1 provides examples of articles that fulfil the conditions previously presented, and, hence, were included in the sample.

The full dataset covers a total of 7,794,312 unique articles, 57,019 of those are associated with at least one keyword from the list, and thus, were included in the calculation of our technology news index (TNI).

However, this method of selecting the relevant sample is obviously not free of drawbacks: 1) potential relevant technological events might not be reported (such as unpatented innovation, as this is a strategic decision); 2) not all technical advances or innovations will materialise into future productivity gains; and 3) as the criteria of the provider or the original source for assigning a given subject keyword to an specific article is not public, it is possible that the final sample include irrelevant texts, which in the end,

[2] `https://developer.nytimes.com/`

Table 1: Examples of articles used in the TNI calculation

Publication Date	Lead Paragraph	Subject Keywords
December 6, 1957	"A system for ordering groceries by telephone, but on a scale undreamed of by the housewife, was patented this week. In a scientific trice, it can transmit to a warehouse orders for thirty-seven cases of eggs, ninety-six bunches of bananas and a half-ton of instant coffee."	- Inventions and Inventors - New US Patents - Food and Grocery Trade
June 3, 1973	"The techniques of manned space flight are finally coming down to earth. The concepts of energy conservation implicit in the lunar excursion module and the space laboratory-the ̈homesöf the astronauts for extended periods of travel in space-are now being applied with dramatic results to the construction and modification of private housing and industrial and commercial buildings. The result has been a saving of millions of dollars a year."	- Energy and Power - Aerospace Industries and Sciences
September 28, 1993	"A CONSORTIUM of American companies and the nation's main nuclear weapons laboratory have joined forces to develop and market a revolutionary new type of computer and television display screen based on synthetic diamond film. The new display screens would exploit a recent discovery that synthetic diamond emits showers of electrons when exposed to weak electric fields. The consortium hopes to steal a long march on competitors in the electronics industry, particularly the Japanese, in a technology expected to earn some $9.4 billion by the year 2000."	- Joint Ventures and Consortiums - Electronics - Television - Diamonds - New Models, Design and Products - Data Processing (Computers)
April 15, 2003	"The human genome is complete and the Human Genome Project is over, leaders of a public consortium of academic centers said today. "We have before us the instruction set that carries each of us from the one-cell egg through adulthood to the grave," Dr. Robert Waterston, a leading genome sequencer, said at a news conference here at the National Institutes of Health."	- Human Genome Project - Genetics and Heredity - Finances - Research - Budgets and Budgeting - DNA (Deoxyribonucleic Acid)

would make the index a noisy measure of true technology news. On the other hand, as pointed out by Barberá et al. (2016), if an strategy that excludes many relevant stories is adopted, this would mean an even noisier measure.

Agents receive a signal from the media coverage through the tone and volume of the economic reporting. In fact, using news corpora from New York Times and Washington Post, Soroka et al. (2015) confirm the hypothesis made by Larsen and Thorsrud (2019b) that changes in the volume of economic reporting are strongly related to current, and, especially, to future changes in economic developments, rather than to past events.

Some papers in the economics literature have explored this characteristic of newspapers and, as in this paper, have relied on frequency-based measures extracted from newspaper corpus for the economic analysis of the role of some structural shocks at micro and macro levels: Alexopoulos and Cohen (2015), Baker et al. (2016) for uncertainty indexes and shocks, and Larsen and Thorsrud (2019a,b) for news shocks.

Following the approach used by these papers, we propose a measure of the intensity in which technology-related articles are published in the The New York Times, the technology news index TNI. Its calculation is straightforward: the total number of articles that are associated with one or more keywords from the list published in a given quarter divided the total number of articles published in that same given quarter and is given by Equation 23.

$$TNI_t = \frac{\#articles\_keywords_t}{\#articles\_published_t} \tag{23}$$

However, due to lack of references in the literature, we hypothesize that the same logic for economic reporting applies for technology news. Hence, intensity in which tech and science news are being published reflects the current efforts, investments and their outcomes in the science, technology and innovation fields, which, as shown, takes some time to be reflected in actual gains in productivity.

Therefore, our index is expected to mainly reflect part of the component “technical change” when it comes to productivity decomposition and its sources of growth. Associated with expansions in the economy’s production possibility set, this component is, according to Fare et al. (1994), a more important driver of productivity gains in developed countries – like the US, which tends to work on the global production frontier – than gains associated with capital deepening or scale and efficiency, which reflect movements towards the frontier.

All in all, we expect that shocks to TNI would lead to a delayed and permanent effect on productivity, resulting in a S-shaped impulse response of TFP, in line with the diffusion of innovations view.

The final TNI is shown in Figure 1. It is possible to associate the spike in 1957 with the beginning of the Space Race. The remarkable growth of the index during 1996-2004 is likely reflecting developments in information technology. For the SVAR estimation, the TNI is used in log-levels.

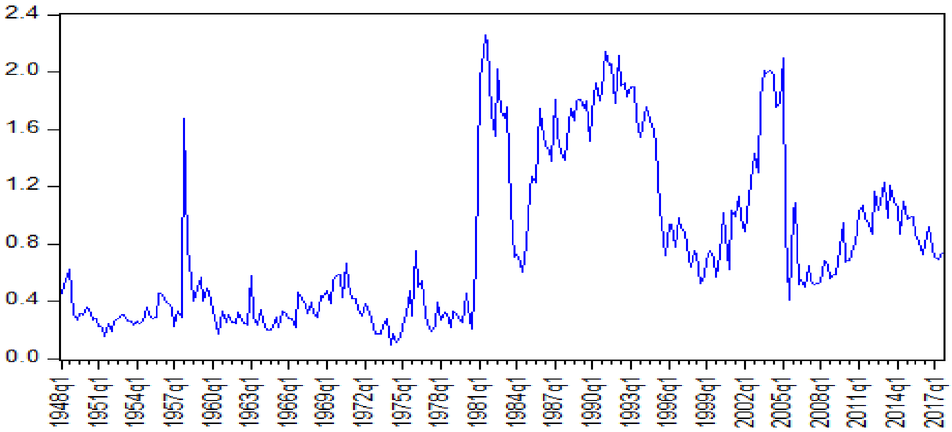

Figure 1: Technology news index (TNI) time-series

The index is a noisy signal of technology news, however it intends to capture the information about technology that are available for the agents, reflecting the signal they recognize and visualize in order to make consumption and investment decisions. Moreover, as TNI is a noisy signal of news, we believe that our results represent a lower bound of the importance of technological news shocks given that it cannot capture all the technical developments that might affect future TFP.

4.2 Other macroeconomic variables

Total factor productivity (TFP) corrected for capacity utilization is originally presented in annualized quarterly growth rates, i.e., $400 * \Delta \log(TFP_t)$, calculated following Fernald (2014)'s methodology, and available at the Federal Reserve Bank of San Francisco website[3]. For this paper, in order to get log figures of TFP, the original data is divided by 400 and then the cumulative sum is calculated for each quarter.

Consumption per capita (C) and investment per capita (I) are quarterly, seasonally adjusted and deflated series supplied by Federal Reserve Bank of St. Louis FRED database[4]. The former is calculated as the logarithm of the sum of personal consumption expenditures of nondurable goods (PCND) and services (PCESV), deflated by each corresponded implicit price deflator (DNDGRD3Q086SBEA and DSERRD3Q086SBEA in FRED, respectively). The latter is the logarithm of the sum of fixed private investment (FPI) and personal consumption expenditures in durable goods (PCDG) deflated by each corresponded implicit price deflator (A007RD3Q086SBEA and DDURRD3Q086SBEA in FRED, respectively). Both per capita aggregates are obtained by dividing the variables by

[3] `https://www.frbsf.org/economic-research/indicators-data/total-factor-productivity-tfp/`
[4] `https://fred.stlouisfed.org`

the civilian noninstitutional population aged from 16 years up (CNP16OV, also retrieved from FRED).

Hours worked (H) is the logarithm of the quarterly annualized seasonally adjusted total hours worked in the US for nonfarm business sectors made available by the US Bureau of Labor Statistics, divided by population CNP16OV.

Labour productivity (L), retrieved from the US Bureau of Labor Statistics (BLS) website, is the logarithm of the quarterly nonfarm business sectors labour productivity index for all employed persons, with 2012=100, and reflects changes in the ratio of output to hours of labour input.

Real gross domestic product per capita (Y), henceforth output, is calculated as the seasonally adjusted real gross domestic product (code GDPC1 from FRED) divided by population CNP16OV.

Figure 2 presents the evolution of these time series in the sample. The figure shows that TFP presents four distinct period of growth and tracks fairly well the developments in technology and productivity for the period studied. Firstly, the Post-World War II period, from 1948 to 1972, with important technological developments due to Astronautics, Nuclear Research and Transistors.

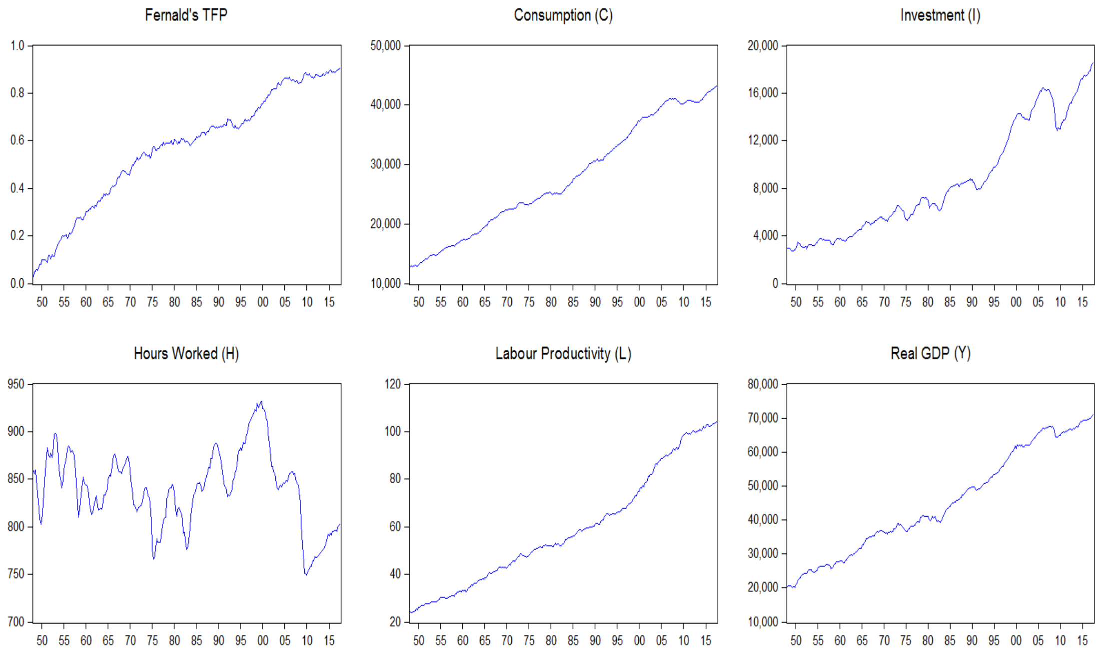

Figure 2: Time series of macroeconomic variables

Second is characterized by a period of low growth due to a 'long slump' period after the 1973 oil crisis from 1973 to 1995. It is then followed by a boom led by the IT revolution, guided by the falling prices in IT and semiconductors from 1996-2004. See Nordhaus (2004), Gordon (2013) and Jorgenson et al. (2014) for a deeper discussion of the historical periods.

5 Results

5.1 Preliminary analysis of the news shock

As a preliminary exercise to understand the data, the cross-correlation coefficients between the constructed TNI and the macroeconomic aggregates of interest are depicted in Figure 3. The figure shows that, the aggregate news index does not seem to have any significant and consistent leading properties with respect to the other variables, suggesting that, at first, TNI does not seem to have any kind of direct relation with them.

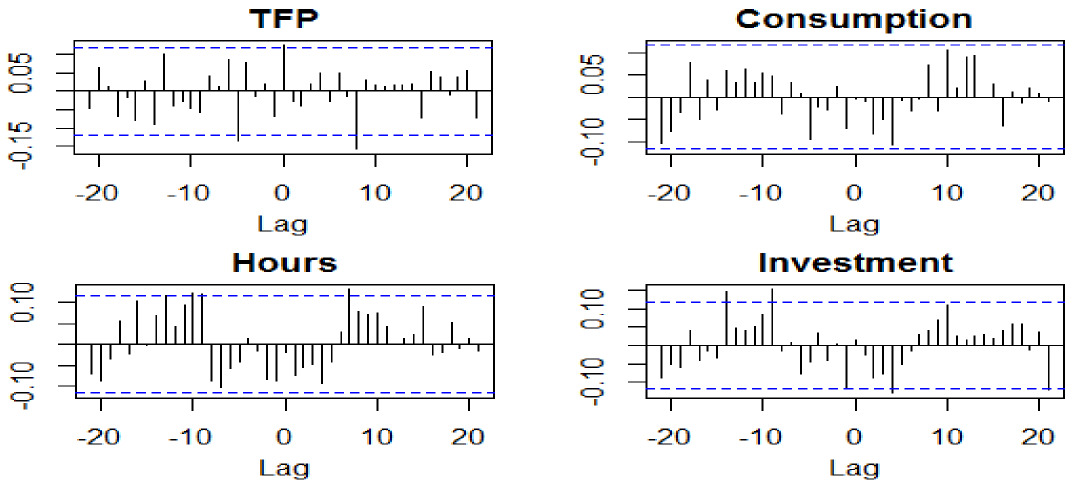

Figure 3: The graph reports the quarter-over-quarter growth rate cross-correlation coefficients between the constructed aggregate technology news index and TFP, consumption, hours worked, and investment, at leads and lags.

Notes: The Y-axis is the cross-correlation coefficients. For negative (positive) values of the X-axis, TNI is leading (lagging) relative to the variables. The broken blue line represents the level in which correlations above it are significant ($\alpha = 0.05$).

Granger Causality Test with four lags cannot reject the hypothesis that TNI Granger-causes ($p - value = 0.0592$) annual TFP growth. Values of TNI throughout the sample provide useful information about future values of TFP growth, whereas, the opposite is not true, it can be rejected that annual TFP growth values Grange-cause TNI.

Next, in order to check whether innovations to TNI are consistent with the expected behaviour of a technology news shock, we estimate a trivariate VAR with TFP as the measure of productivity and consumption which is a forward-looking variable that captures the news shock, so that $y_t = [TFP_t \; TNI_t \; C_t]$. According to the theoretical model it can immediately react to changes in information.

We estimate the structural vector autoregressive model using quarterly data for the full coverage of the news from 1948Q1 to 2017Q3. The identification scheme combines short and long-run restrictions, where TFP and TNI are ordered first in the system, hence news shock is identified as the innovation to TNI that is orthogonal to TFP. The approach follows closely the paper by Beaudry and Portier (2006). The model is estimated in levels and the lag order is assumed to be four. Subsection 5.5 discusses robustness checks to different number of lags among other changes to the model.

Given the description of the identification strategy for the trivariate system exposed in the Subsection 3.2, the news shock ϵ_2 is identified as the innovation to TNI that does not affect TFP on impact, but does on the long run, and is equivalent to η_2 in Equation 16. Additionally, such identification scheme allows for a traditional supply shock (unanticipated technology shock), ϵ_1, an innovation to TFP, fully unrestricted, and a traditional demand shock ϵ_3, an innovation to Consumption (C) with temporary effects on both productivity and consumption.

Figure 4 depicts the impulse response function (IRF) graphs following a news shock, ϵ_2. It is possible to see in the figure that ϵ_2 carries information about future productivity growth: TFP takes some time to respond positively to an initial innovation in TNI - approximately 13 quarters - which suggests that technical changes are not instantaneously reflected in productivity gains. The shock causes an S-shaped curve. The diffusion from innovation to productivity is initially rather slow, followed by a period of rapid increase and stabilises when it is fully absorbed in the economy, as documented, e.g., in Rogers (2003).

The short-run negative response in TFP may be caused by an excessive adjustment of capital utilisation, according to Beaudry and Portier (2006), or due to labour reallocation towards the new technology associated with a learning process, as documented in Christiansen (2008). Permanent changes in TFP is preceded by an increase in TNI.

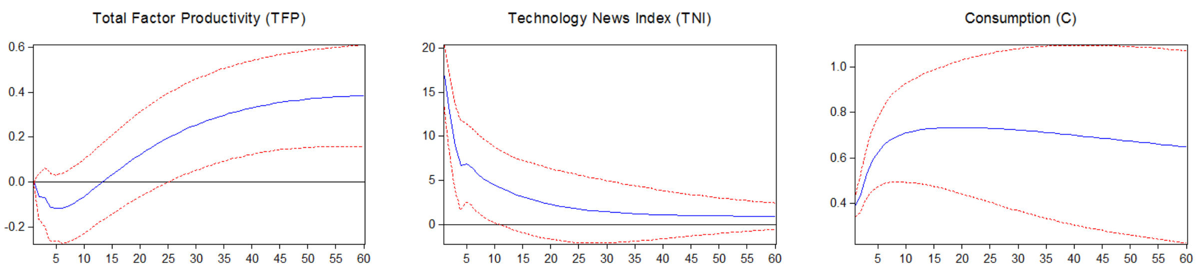

Figure 4: Impulse responses to a news shock, ϵ_2, in the trivariate system $[TFP_t \; TNI_t \; C_t]$.
Notes: one-standard deviation ϵ_2 shock in percent deviation from the steady state. The blue solid line represents the point estimate and red dashed lines represent 95 percent confidence bands.

The response of consumption to the arrival of news about future TFP is positive already on impact, before any actual gains in productivity. This result is consistent with the theoretical model presented in the Section 3. News create a positive wealth shock. The expectation about higher income in the future induces an increase in consumption today, which is in line with both dominant macroeconomic views. However, by the time the TFP starts to increase, the dynamics of consumption seems to be exhausted, returning very slowly back to the trend, though, confidence intervals remain rather wide.

To sum up, the evidence is consistent with the view that agents update their economic perception, and are able to anticipate future productivity due to the lagged effect that technology news have in TFP. Differently to a noise shock, the findings here are in line with those predicted by the model described in Subsection 3.1, in which a news shock have a delayed and permanent effect on productivity, and affects consumption on impact and permanently.

It also worth analysing the responses following a surprise technology shock, ϵ_1, and a demand shock, ϵ_3. As seen in Figure 5 following an unanticipated technology shock both TFP and consumption rise on impact, but they both tend to return to their steady state quickly, hence, although not restricted to have long run impact on TFP, ϵ_1 acts like a temporary TFP shock, while ϵ_2 is the shock that permanently affect productivity, and is orthogonal to current TFP. ϵ_3 only have temporary effects on both TFP (decreases on impact, probably reflecting the intertemporal substitution between productivity-improving and directly productive activities throughout business cycles, as reported by Saint-Paul (1993)) and Consumption (increases on impact, as expected from a demand shock).

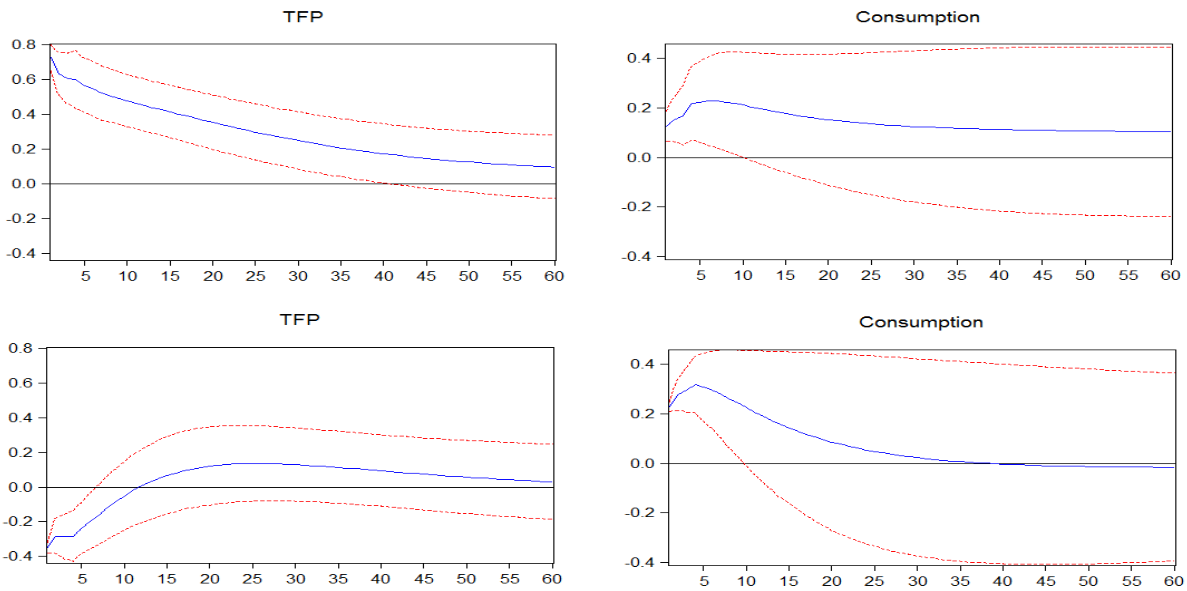

Figure 5: Impulse responses to an unanticipated technology shock, ϵ_1 (upper panels), and a demand shock, ϵ_3 (lower panels) in the trivariate system $[TFP_t\ TNI_t\ C_t]$.

Notes: one-standard deviation ϵ_1 and ϵ_3 shock in percent deviation from the steady state. The blue solid line represents the point estimate and red dashed lines represent 95 percent confidence bands.

The results presented in this section casts doubt on the relevance of traditional technology shocks as important drivers of business cycles, whereas, it does provide support for the news-driven business cycles.

5.2 Main results

The main objective of the paper is to study macroeconomics fluctuations, so we augment the trivariate specification with additional variables of central interest in business cycles analysis. Firstly, hours worked (H) is added to the system (the baseline model from now on) in order to analyse whether news shocks can cause positive comovements between consumption and hours, as claimed by the news-driven business cycles view, or it causes recessions or opposite movements, as predicted by the permanent income hypothesis.

The identification scheme follows very closely the approach in the previous system, except for the fact that now ϵ_4, an hours-specific shock is present, meaning that it is allowed to affect only H on impact, which can be interpreted as a measurement error in hours worked, as suggested by Beaudry and Portier (2006).

Figure 6 reports the impulse responses graphs of TFP, consumption and hours to a technology news shock ϵ_2. Over the first ten quarters, the dynamics of the responses is very rich: both TFP and consumption follow the same pattern as in the trivariate case. Consumption increases by 0.4 percent on impact and reaches the peak of about 0.7 percent by the 6-quarter.

The impact response of hours is 0.24 percent. A temporary boom follows so that the dynamics reaches a peak of 1.02 by the 6th quarter after the initial shock. After the hours start reverting back to the steady state, resulting in a hump-shaped response. Such dynamics of the temporary boom can be explained as being a period when agents make decisions in order to take advantages of future productivity developments due to arrival of news (Beaudry and Portier, 2006).

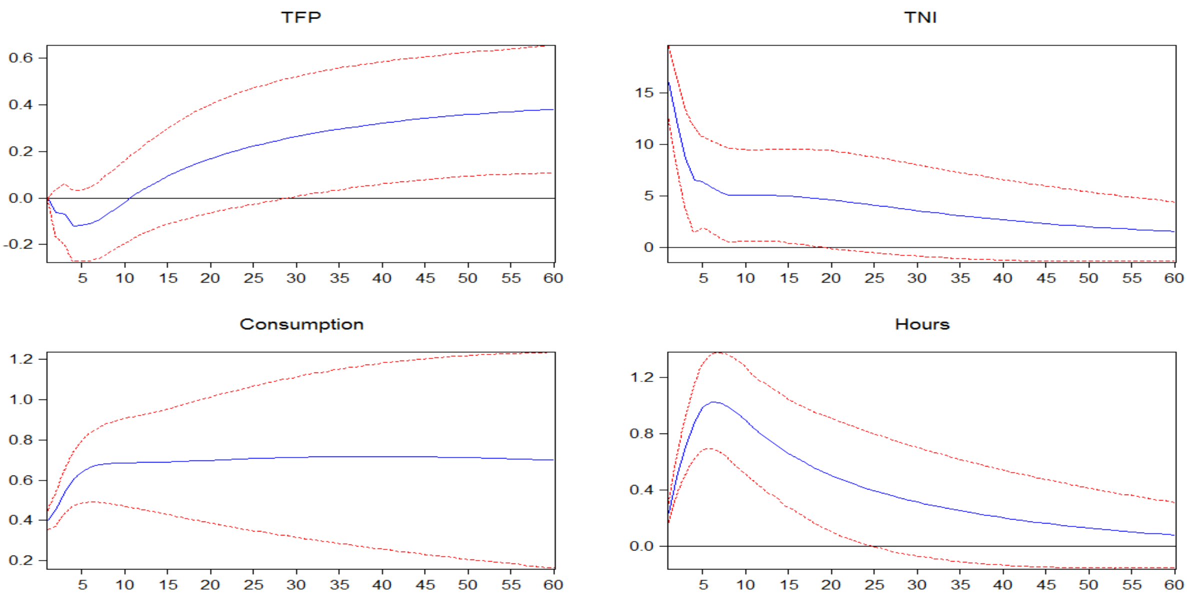

Figure 6: Impulse responses to a news shock, ϵ_2, in the 4-variable system $[TFP_t\ TNI_t\ C_t\ H_t]$.

Notes: one-standard deviation ϵ_2 shock in percent deviation from the steady state. The blue solid line represents the point estimate and red dashed lines represent 95 percent confidence bands.

The reasons behind the initial boom in hours becomes more evident when investment is included in the 4-variable system instead of hours. The figure 7 shows that within the first 7 quarters, as producers see the opportunity in the economy, they create an investment boom, and, in consequence, due to frictions in the labour market, firms have incentives to start hiring ahead of an anticipated improvement in technology, as in Faccini and Melosi, 2018), leading to a boom in hours worked.

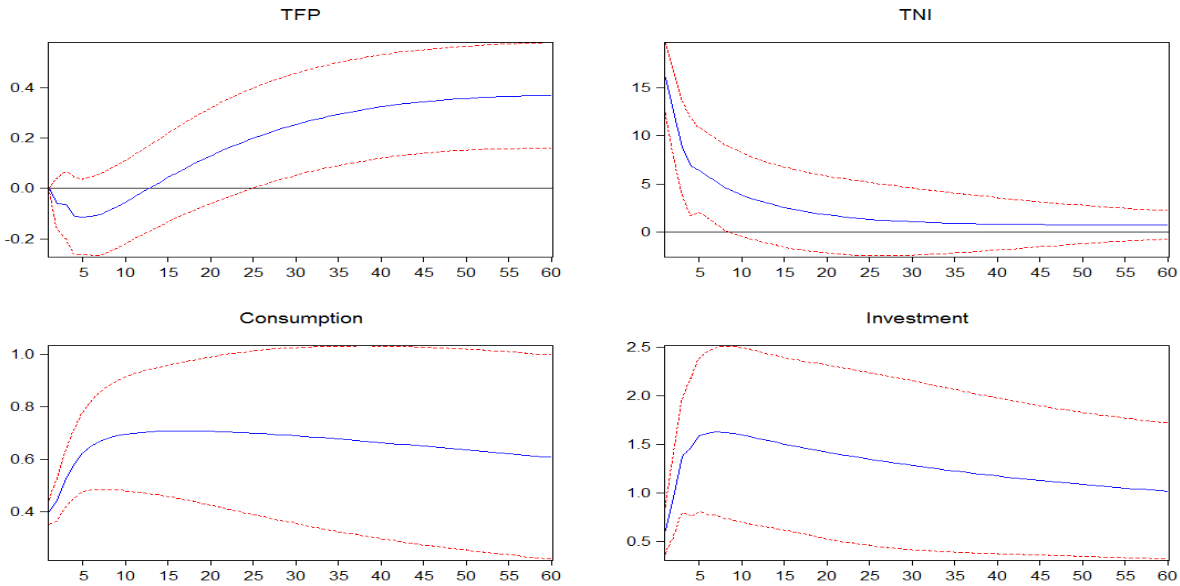

Figure 7: Impulse responses to a news shock, ϵ_2, in the 4-variable system $[TFP_t\ TNI_t\ C_t\ I_t]$.

Notes: one-standard deviation ϵ_2 shock in percent deviation from the steady state. The blue solid line represents the point estimate and red dashed lines represent 95 percent confidence bands.

One of the goals of the paper is to the identify news shocks using the TNI as some noisy measurement of scientific and technological development. As shown before technical change is the main component of productivity growth in developed countries. Thus, we expect that the shock identified here, is able to explain a considerable share of the FEVD in the medium and long-run, while zero or almost zero in the short-run, given the slow diffusion of innovations.

Figure 8 shows that within the first 17 quarters, the relative importance of ϵ_2 is almost negligible when compared to the traditional surprise technology shock ϵ_1, but starts rising gradually, and as the horison tends to the infinity, it becomes more and more important, surpassing ϵ_1 in importance after approximately 81 quarters. Which is to some extent in accordance with the speed of the diffusion discussed in the literature, where it can varies, on average, from 5 to 15 years in which "half of the innovation's adopters do so" (Rotemberg, 2003) or 32 until it reaches the peak of adopters (Gort and Klepper, 1982). These results provide another evidence that TNI, although noisy, carries fundamental information about future productivity with no impact effect, an important characteristic of a typical news shock described in the literature.

When the baseline specification is estimated with output (Y) in place of H, we can

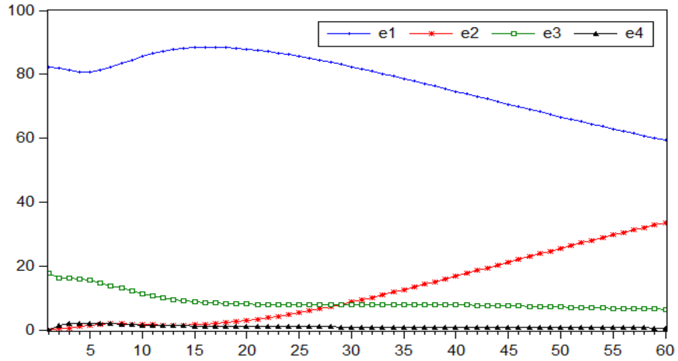

Figure 8: Forecast Error Variance Decomposition of TFP attributable to shocks in the system: the surprise technology shock ϵ_1, the news shock ϵ_2, the demand shock ϵ_3, and an hours-specific shock ϵ_4.

Notes: The FEVD for TFP is from the $[TFP_t\ TNI_t\ C_t\ H_t]$ specification.

check whether the responses are in line with predictions of the model presented in Subsection 3.1. As we stated that our TNI is supposed to be a noisy signal of news, we expect that both consumption and output respond positively on impact, present similar dynamics, and, given the informational and nominal rigidities in the economy, peak quarters after the news arrive before relevant gains in TFP. Figure 9 shows that the dynamics of the economy is in accordance with a model of news-driven business cycles. C jumps immediately by 0.4% and exhausts its dynamics after about 6 quarters, while Y increases by 0.3% on impact and reaches its peak of 0.8% after 5 quarters, with both remaining at a higher new level for a long period. While, TFP only becomes positive after 12 quarters.

Similarly to the the analysis made in the previous subsection, Figure 10 depicts the responses following ϵ_1 and ϵ_3 shocks. From the figure is possible to see that the traditional surprise technology shock is not able to produce the comovements expected in business cycles, although it positively affects consumption on impact, it leads to a decrease in hours worked (H), in accordance with Smets and Wouters (2007). On the other hand, demand shocks produces comovements, but the responses are only temporary. The results suggests then that news shocks are a very interesting component, that behaves like demand shocks in the short and medium-run, and as a supply shock in the long run.

Now that the evidences suggest that news shocks are able to produce comovements between macroeconomic variables, we are also interested in determining the importance of technology news shocks in driving business cycles. For this purpose, Figure 11 presents the Forecast Error Variance Decomposition (FEVD) of important macroeconomic variables to the technology news shock, ϵ_2, retrieved from: the baseline four-variable model; from a model with investment (I) in place of hours worked (H) (see Figure 7 for the impulse responses); and from a model in which H is replaced by Y, which impulse responses are depicted in the Figure 9.

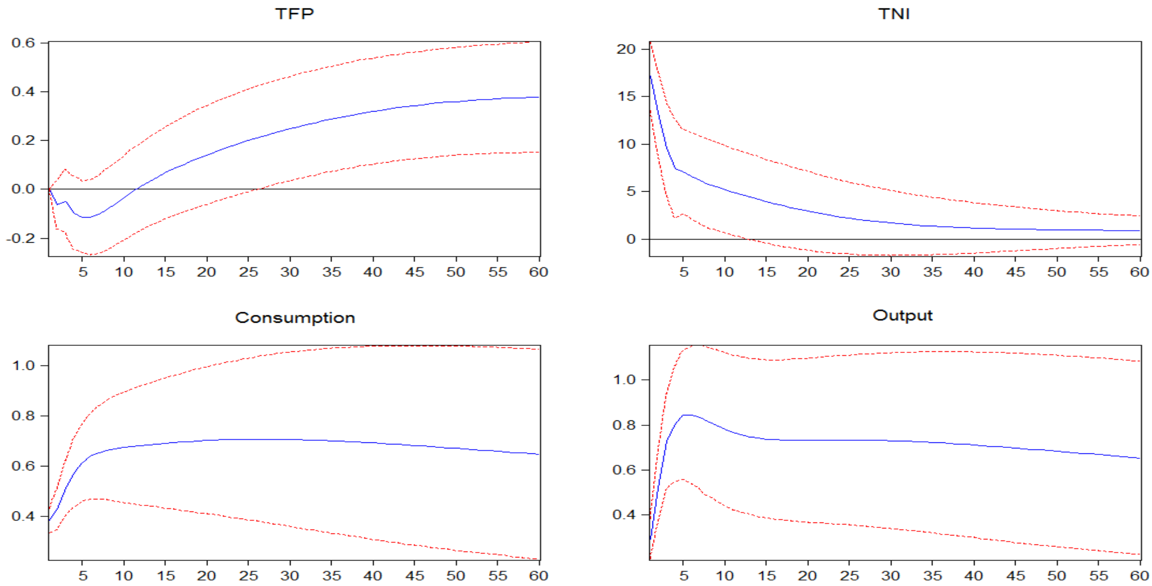

Figure 9: Impulse responses to a news shock, ϵ_2, with output (Y) in place of hours worked (H) in the baseline specification.

Notes: one-standard deviation ϵ_2 shock in percent deviation from the steady state. The blue solid line represents the point estimate and red dashed lines represent 95 percent confidence bands.

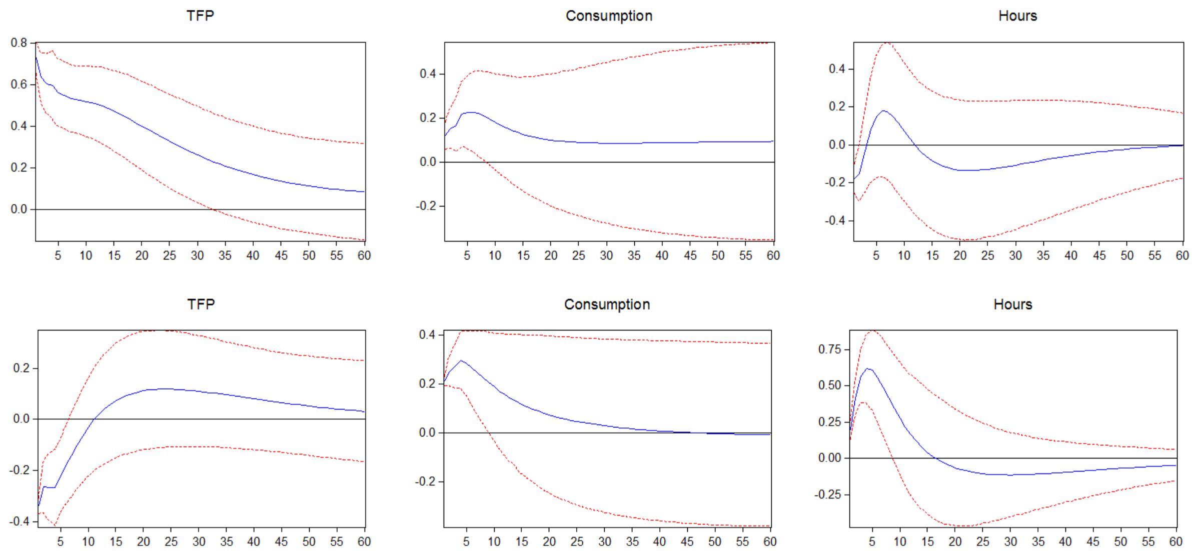

Figure 10: Impulse responses to an unanticipated technology shock, ϵ_1 (upper panels), and a demand shock, ϵ_3 (lower panels), in the four-variable system $[TFP_t \ TNI_t \ C_t \ H_t]$.

Notes: The graphs in the figure depicts the impulse response to a one-standard deviation ϵ_1 and ϵ_3 shocks in percent deviation from the steady state. The blue line represents the point estimate of the Impulse Response Functions. The red dotted lines represent 95-percent confidence bands.

The left-hand panel of Figure 11 shows that the effects of news shocks in the business cycles is rather important, accounting for at most 67% of the output forecast error variance at business cycles frequency (8 to 32 quarters), suggesting that, indeed, shocks that reflect new about future productivity are a major source of macroeconomic fluctuations, consistent with the news-driven business cycles view. Moreover, the figure indicates that these shocks explain most of the variance of consumption and a relevant share of hours and investment forecast error variance at business cycles frequencies, while accounting for a small share of TFP at short and medium frequencies. Whereas news shocks are central to business cycle fluctuations, a surprise technology shock, ϵ_1, are not as important, as seen in the right-hand panel of the Figure 11

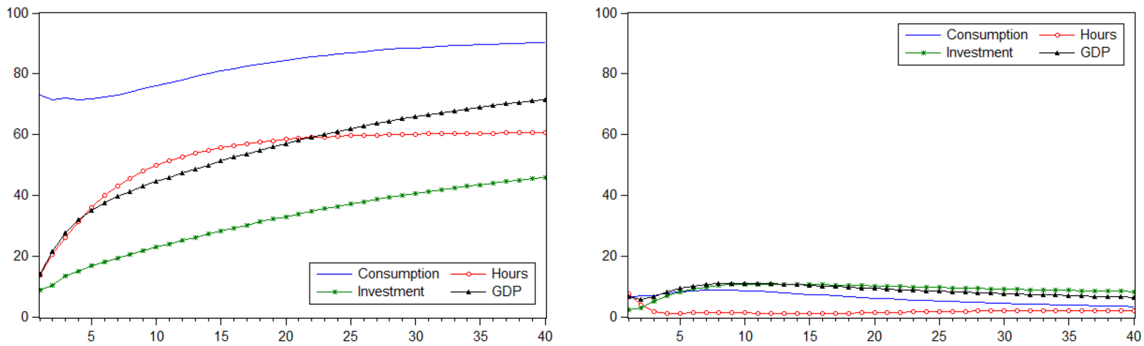

Figure 11: Forecast error variance decomposition of consumption, hours worked, investment and output attributable to a one standard deviation ϵ_2 news shock (left panel) and to a one standard deviation ϵ_1 surprise technology shock (right panel).

Notes: The FEVD for consumption and investment is from the $[TFP_t \ TNI_t \ C_t \ I_t]$ specification, for hours worked is estimated from the $[TFP_t \ TNI_t \ C_t \ H_t]$ system, and for Y from the four-variable SVAR specification $[TFP_t \ TNI_t \ C_t \ Y]$.

Additionally, Appendix B.1 provides additional results using a sentiment-based index constructed from the selected technology articles. Although sentiment content might provide a hints towards future developments of productivity, the SVAR estimation with this variable did not present a good robustness to different specifications compared to the intensity-based one used in the benchmark model in this paper.

5.3 News and long-run developments in productivity

In this section, it is shown that the identified shock ϵ_2 is almost perfectly collinear to a shock $\tilde{\epsilon_1}$ that drives long-run movements in TFP.

One of the predictions of the model with delayed response of innovation on productivity proposed by Beaudry and Portier (2006) is that in a VAR specification where TFP and Stock Prices are placed first in the system, the shock ϵ_2 identified with a combination of short and long-run specification will be almost perfectly colinear and produce the same dynamics as a shock $\tilde{\epsilon_1}$ identified with recursive long-run restrictions, which is equivalent to shock that drives the TFP path in the long-run.

The hypothesis above is tested using the TNI instead of Stock Prices in the benchmark 4-variable specification SVAR estimated with four lags, with the system $[TFP_t\ TNI_t\ C_t\ H_t]$.

Figure 12 shows that the dynamics created by the shock $\tilde{\epsilon_1}$ is very similar to the impulse responses following the ϵ_2 shock, even though, both shocks were supposed to be completely different (orthogonal), as $\tilde{\epsilon_1}$ is an innovation to TFP, and ϵ_2 is an innovation to TNI that is contemporaneously orthogonal to TFP.

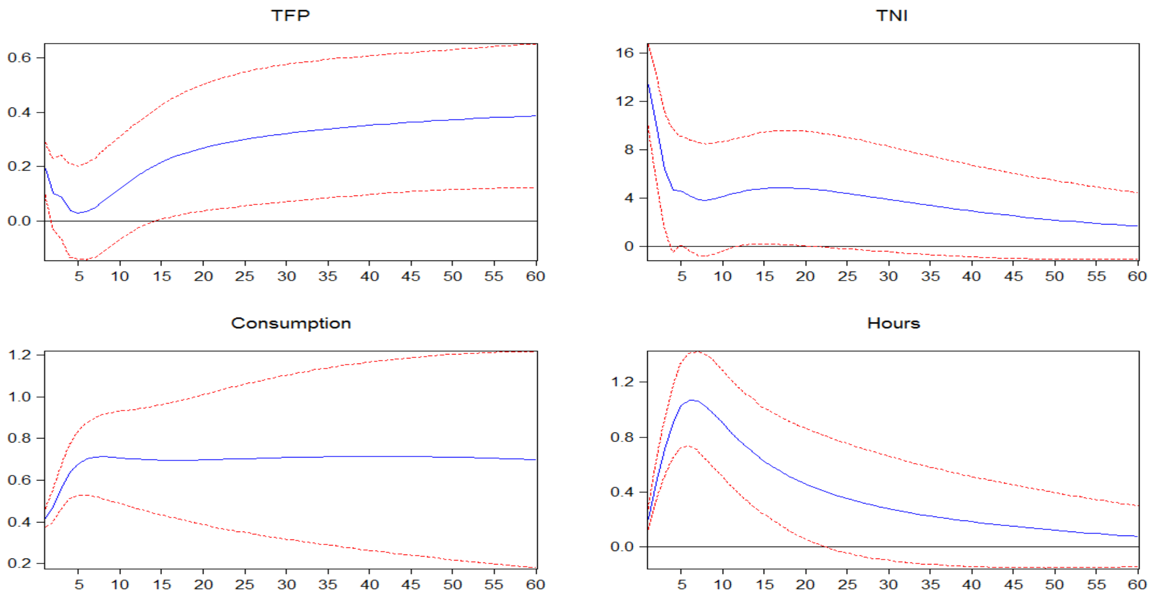

Figure 12: Impulse responses to a technology shock, $\tilde{\epsilon_1}$, in the 4-variable system $[TFP_t\ TNI_t\ C_t\ H_t]$ with recursive long-run identification.

Notes: Each graph in the figure depicts the impulse response to a one-standard deviation $\tilde{\epsilon_1}$ shock in percent deviation from the steady state. The blue line represents the point estimate of the Impulse Response Functions. The red dotted lines represent 95-percent confidence bands.

In fact, when Figure 12 is compared to Figure 6, it is possible to see that $\tilde{\epsilon_1}$, as expected, has a permanent effect on TFP, but, more surprisingly, the impact effect on TFP is almost null, while it has a considerable impact effect on TNI. It reinforces the idea that TNI is indeed able to identify a technology shock that have delayed effects on TFP, and that, permanent and long-run developments in TFP are reflected in the TNI before they effectively shift outwards the production frontier. Furthermore, there is a lag between the time technology news are published (and observed by agents) and the time when they are reflected in gains of productivity, as expected for a news shock and by Beaudry and Portier (2006) model.

Moreover, Figure 13 depicts the time series of ϵ_2 and $\tilde{\epsilon_1}$ in the left panel, and the scatter plot of ϵ_2 against $\tilde{\epsilon_1}$ in the right panel. The figure shows that both shocks are highly correlated, which clearly supports the idea that both orthogonalization schemes recover essentially the same shocks.

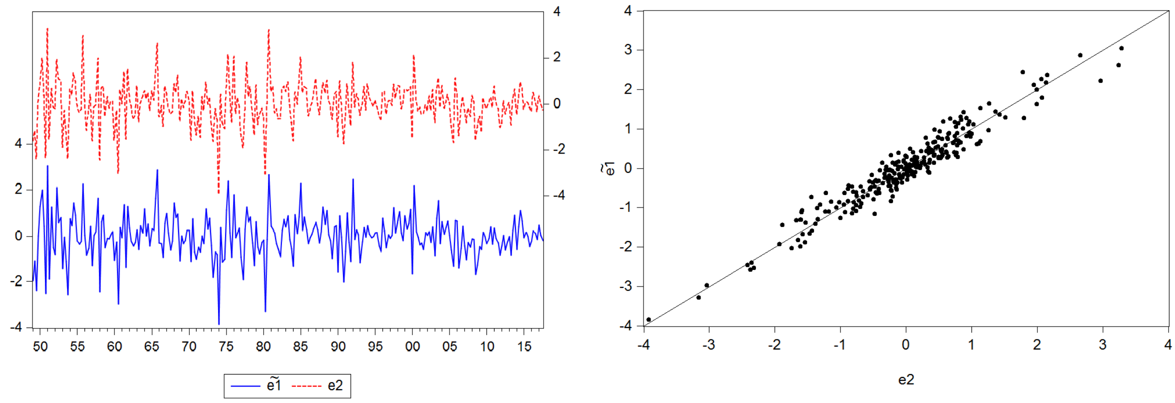

Figure 13: Time series of ϵ_2 and $\tilde{\epsilon_1}$ (left panel). Scatter plot of ϵ_2 against $\tilde{\epsilon_1}$ (right panel).
Notes: The straight line in the scatter plot is a 45-degree line. Both shocks are obtained from the benchmark 4-variable specification, but with two different identification schemes. Correlation coefficient between ϵ_2 and $\tilde{\epsilon_1}$ is 0.963.

5.4 Discussion

Many papers used different techniques for the identification of news shocks under SVAR frameworks, some researchers identified them as innovations to forward-looking variables, some used innovations to more direct measurements of technology, such as patents, and, others suggested the identification through schemes that rely on the maximum share of explained TFP forecast error variance at/over one or some horizons. In this section we place our findings in the context of this existing empirical literature, and evaluate our results against the different approaches proposed in the news shocks literature.

First, we discuss how the literature analyses news shocks propagation into productivity, and its importance for the forecast error variance of TFP in the short and the long-run.

Then, we discuss in which context macroeconomic comovements arise in the economy, compare how different methods evaluate the importance of news shocks for business cycles fluctuations, and place our results in the existing literature.

News Shocks and Total Factor Productivity. Much of the literature has found that news shocks already increases TFP in the short run with permanent effects, however, as pointed out by Barsky et al. (2014) and Portier (2014), for shocks with such dynamics, it is very difficult to identify pure news shocks from the effect of actual TFP, important spillovers might had occured during the transmission of news shocks. Although still predicts future evolution of productivity, the movements of macroeconomic variables at medium frequencies might not be the effect of news per se, but rather the effect of the subsequent actual improvement in technology.

Thim seems to be the case of Barsky and Sims (2011), Barsky et al. (2014) and Forni et al. (2014), that use identification strategies that identify news shocks as the one that maximize TFP variance over/for a given horison. For all these cases, most of TFP adjustments are done in the first quarters. Larsen and Thorsrud (2019b) uses a different approach, but the outcome is similar: TFP peaks very quickly and news shocks do not seem to have a relevant role in the TFP's path in the long-run, although not explicit, its contribution stagnates at 7% from the 20-quarter horison onwards. For these cases, the FEVD of TFP to a news shock shows that the dynamics is already exhausted in the short to medium run, with not much gains in the long run.

The results presented here are less subject to this criticism, because the response of TFP to a news shock resembles those of what Portier (2014) called "Technological Diffusion News", in which news carries fundamental information about future developments of TFP without short run effect on TFP (in fact, up to 2%, 10% and 60% at short, medium and long frequencies, respectively, of TFP's forecast error variance is due to TNI news shocks), and are more likely to produce the typical comovements of business cycles, as it is going to be shown in the next section. This S-shaped response of TFP seems to be a common feature of news shocks identified in a Beaudry and Portier (2006)'s fashion, combining short and long-run restrictions, e.g., the present paper, Beaudry and Portier (2006) itself, Beaudry and Lucke (2010), and Beaudry and Portier (2014), and also a feature of news shocks identified under a SVAR-IV framework, as in Miranda-Agrippino et al. (2018) and Cascaldi-Garcia (2019).

News Shocks and Business Cycles. In the core of the news-driven business cycles literature is the analysis of news shocks causing positive comovements between macroeconomic variables, and if they do, to what extent they are relevant at business cycles frequencies (2-8 years).

Some of the papers from the previous section have found that the wealth effect is higher than the substitution effect following a news shock, in favour of traditional RBC models. Among them, Barsky et al. (2014) found that although consumption, investment and output comoves positively, hours decrease on impact. Barsky and Sims (2011) shows that, following a positive news shock, consumption increases, while output, hours, and investment decreases on impact, though this shock is a relevant source of output fluctuations at medium frequencies. In Forni et al. (2014), investment and output fall on impact and then gradually reach their new long-run level; consumption is not affect on impact, but starts increasing after one quarter; while hours decline in the short run. Interestingly, these cases share the similar max-share identification strategy, and innovations diffuses very rapidly.

Larsen and Thorsrud (2019b) is a case in which, as analysed before, TFP peaks very quickly after their identified news shock, while consumption, employment, hours worked, investment and output positively comove. However, the dynamics of these variables seem to track, rather than anticipate TFP developments. They report that news shocks can account for up to 17%, 20% and 8% of the variances of consumption, employment and hours worked, respectively, at business cycles frequencies, but, given the nature of the response of TFP, the FEVD of these variables might be confounding effects that are not caused by news shocks per se, but for actual gains in productivity.

In the last category, news shocks produce macroeconomic comovements and play a relevant role as drivers of business cycles fluctuations. Not by coincidence (see Portier (2014) for discussion on technological diffusion news and business cycles), papers in this category were able to obtain a slow diffusion response of TFP to news. This is the case of papers mentioned in the last section that either applied similar identification to Beaudry and Portier (2006), or the SVAR-IV identification. In all theses cases, the contribution of news shocks to TFP variance in the short and medium-run is very small, and their maximum contribution of news shocks for output fluctuations at business cycles are non-negligible, ranging from about 14% in Miranda-Agrippino et al. (2018) to 82% in the trivariate level specification of Beaudry and Portier (2014), consistent with what is expected from an "almost pure" news shock.

Results of this paper confirm the findings of this category, in which technology news shocks resemble the slow diffusion of innovations and can account for a significant share of business cycles fluctuations, up to 67%, and the long-run TFP's variance (61% at the 120-quarter horison). Looking at Figures 8 and 11, it is possible to see that substantial gains in consumption, investments, hours and output occurs quarters before news shocks actually have relevant impact on TFP (which occurs after about 16 quarters), suggesting that, indeed, much of the variation of these variables was due to "pure" news shocks.

5.5 Robustness

This section offers a series of robustness tests to the baseline specification [TFP_t TNI_t C_t H_t]. First, the robustness is tested for the lag length selection, next for the pre-Great Recession sample, for the Beaudry and Portier (2006)'s original dataset, for an alternative measure of productivity, namely labour productivity, that compares the economy's output with the number of hours worked to produce this output, and, finally, robustness checks regarding the structural break in TNI between 1980Q4 and 1981Q1 are performed.

The lag length selected for the baseline model in the paper was four given the quarterly frequency of the dataset. However, we show that the results are robust independently of the number of lags selected as they do not change the results significantly. The impulse response responses associated with each lag length selection are depicted on Figure 14. It is possible to visualize that all models present very similar patterns and magnitudes for the responses to the news shock ϵ_2. The responses in the baseline model are slightly higher in the short-run, while for the model with 6 lags the responses tend to be higher as the horison goes to infinity.

As seen in Figure 2, there was an abrupt change in the path of the macroeconomic variables due to the Great Recession in 2008, so it makes sense to analyse whether the model is robust when considering periods pre-global financial crisis and pre-Zero Lower Bound. In Figure 15 the impulse responses to a news shock ϵ_2 for the pre-crisis sample are

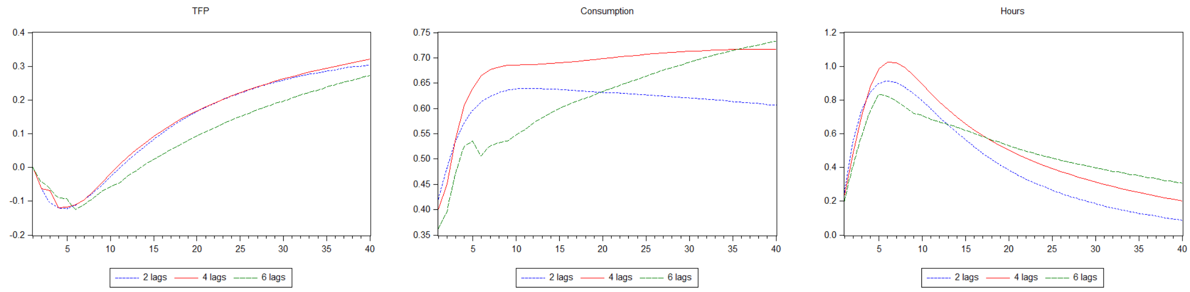

Figure 14: Robustness to lag length selection.

Notes: one-standard deviation news shock ϵ_2 in percent deviation from the steady state. The blue solid line represents the point estimate and red dashed lines represent 95 percent confidence bands.

depicted. The responses are slightly lower and last for less time for the shorter sample, we verify that neither events considerably affect our results.

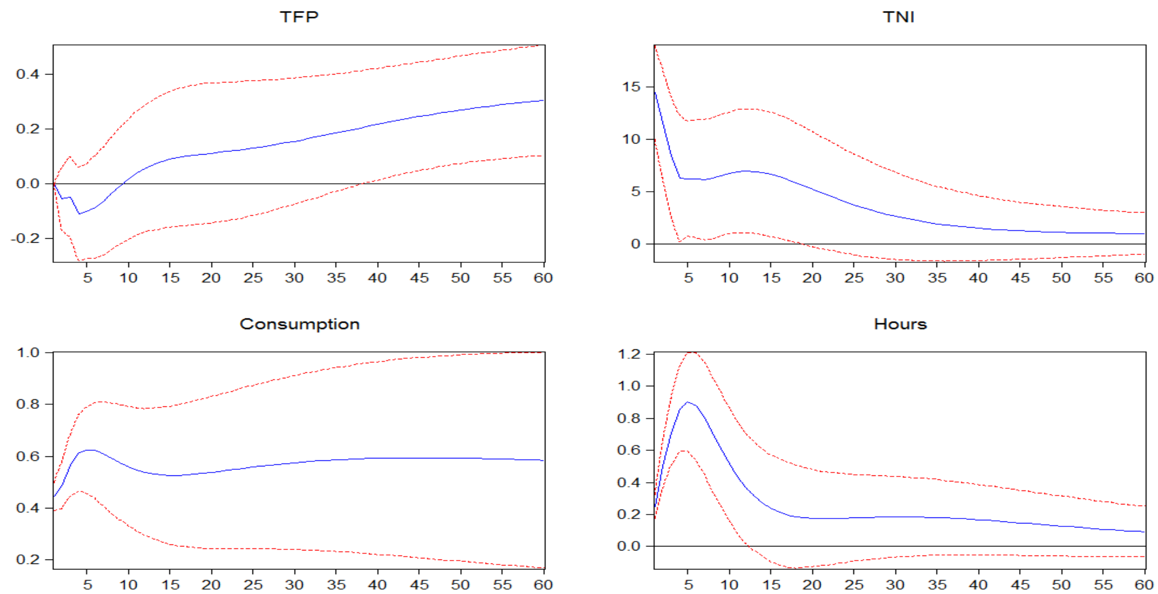

Figure 15: Impulse responses to a news shock, ϵ_2, in the benchmark SVAR estimation for the pre-Great Recession sample (1948Q1-2007Q4).

Notes: one-standard deviation ϵ_2 shock in percent deviation from the steady state. The blue solid line represents the point estimate and red dashed lines represent 95 percent confidence bands.

As the identification scheme used in this paper is the same as in Beaudry and Portier (2006), another valuable robustness exercise is placing our constructed TNI in place of asset prices in the original dataset used by them[5]. One characteristic of their dataset is that they use a different measure of utilization-adjusted TFP, TFP^A, calculated as $TFP_t^A = log(Y_t/H_t^{\bar{s}_h}(CU_tKS_t)^{1-\bar{s}_h})$, where Y is output, H is hours worked, $\bar{s}_h$ is the average level of the labour share over the period, CU is the BLS's measure of capital utilisation and KS is capital services. We estimate the model with the four-variable specification $[TFP_t^A\ TNI_t\ C_t\ H_t]$ and four lags for the time span from 1948Q1 to 2000Q4. Impulse responses

[5]Available at `https://www.aeaweb.org/aer/data/sept06/20030282_data.zip`

to a ϵ_2 shock are shown in Figure 16. The dynamics from both models are the same. So, there are no major changes when we apply our constructed index in the original dataset used by Beaudry and Portier on their influential work.

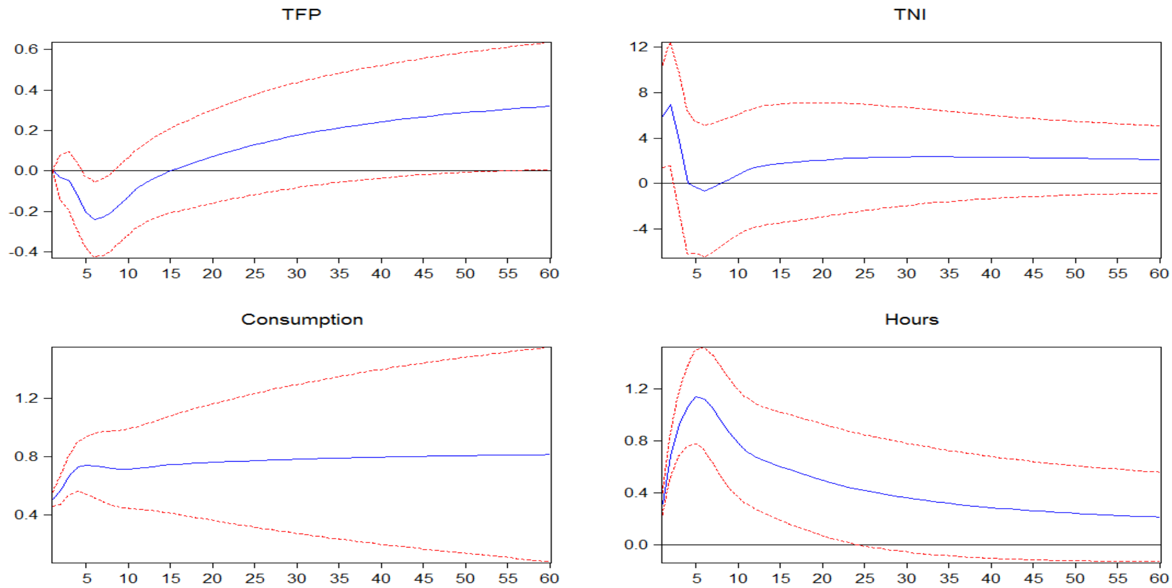

Figure 16: Impulse response to a news shock, ϵ_2, in the benchmark SVAR estimation with Beaudry and Portier (2006)'s original data and TNI for the time span from 1948Q1 to 2000Q4.

Notes: one-standard deviation ϵ_2 shock in percent deviation from the steady state. The blue solid line represents the point estimate and red dashed lines represent 95 percent confidence bands.

Labour productivity (L) is often used as a measure of productivity for international comparison because of its availability across different countries, and due to the limited availability of TFP, especially at quarterly frequency. So, for the sake of future extension of the method presented here to other countries, it is valid to work with labour productivity instead of TFP and analyse whether the results are similar for both cases. And, in fact, this is what happens as shown in Figure 17. However, it is worth noticing that, within the first 10 quarters, the response of labour productivity and TFP are different. Instead of declining and then recovering, like the TFP, there is a temporary boom in L, resembling the response of the unadjusted for capital utilisation measure of TFP used in Beaudry and Portier (2006), which captures the cyclical behaviour of capital utilisation rates, or maybe, simply because the changes in Y are higher than the variations in H. Nevertheless, after 10 quarters, both responses are very similar. However, the model with capital-adjusted TFP is more consistent with what it is expected from the diffusion process of innovations.

When we observe the Figure 1, it is possible to see that there is a break between 1980Q4 and 1981Q1 representing a change from a low to a high state, so in order to test whether this break affects the general results of the paper, we perform two robustness tests.

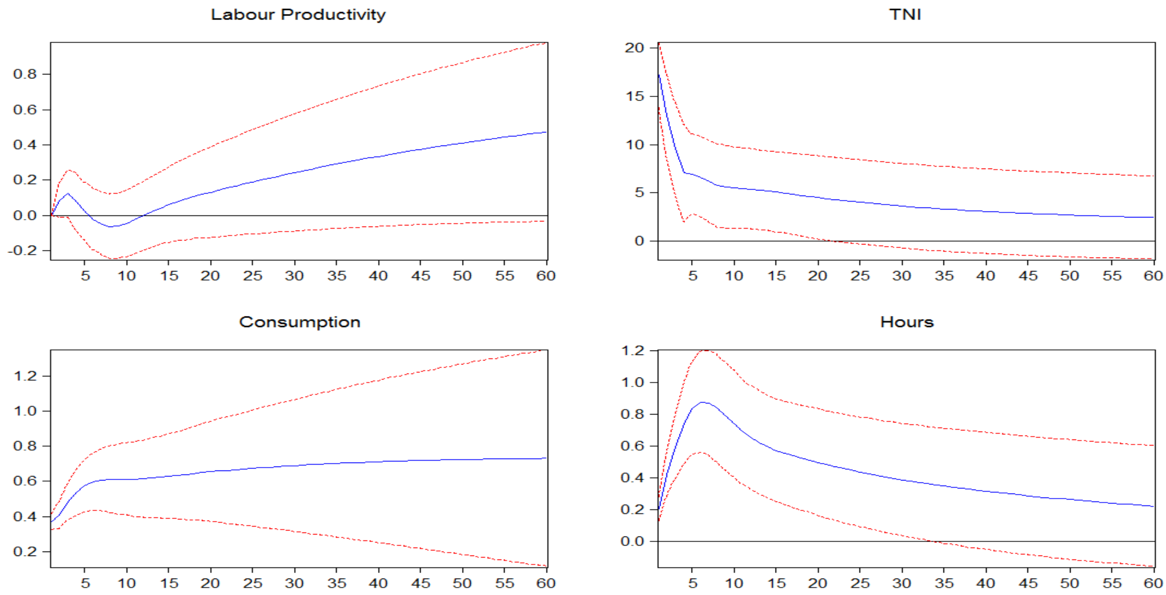

Figure 17: Impulse response to a news shock, ϵ_2, in the benchmark SVAR estimation with labour productivity in place of Fernald's TFP.

Notes: one-standard deviation ϵ_2 shock in percent deviation from the steady state. The blue solid line represents the point estimate and red dashed lines represent 95 percent confidence bands.

For the first test, we split the full sample into a pre-break sample from 1948Q1 to 1980Q4, and a post-break sample from 1981Q1 to 2017Q3, and estimate two VARs with three lags using each sample individually. The results are depicted in Figure 18. It is possible to visualise that for the pre-break sample, TFP responds positively to a news shock quicker than in both the full-sample model, depicted in Figure 6, and the post-break estimate, suggesting that the diffusion of innovations occurred in a faster pace before 1981. However, regardless of the sample used, the dynamics are in line with what is expected following a diffusion news shock, i.e., TFP follows a diffusion process, while consumption and hours reach their peak before relevant actual gains in productivity.

For the second test regarding the break in 1981, we create a dummy variable with values *zero* for the low state, i.e., pre-break, and values *one* for the post-break sample, and estimate the baseline model incorporating the dummy as an exogenous variable. Figure 19 depicts the Impulse Response Functions of the estimated SVAR following a news shock. The figure shows that for the "dummy VAR" estimation, even though TFP turns positive earlier, the pattern and dynamics are essentially the same as for the baseline model, with no relevant implications for the general results.

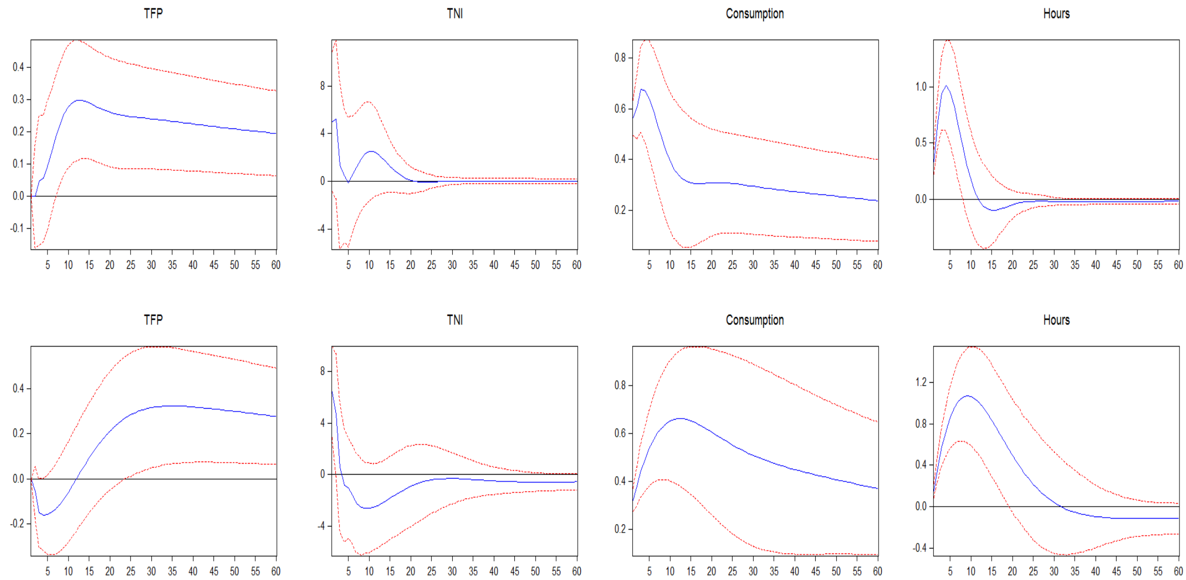

Figure 18: Impulse response to a news shock, ϵ_2, in the benchmark SVAR estimation for the pre-break sample, from 1948Q1 to 1980Q4 (upper panels) and for the post-break sample, from 1981Q1 to 2017Q3 (lower panels).

Notes: one-standard deviation ϵ_2 shock in percent deviation from the steady state. The blue solid line represents the point estimate and red dashed lines represent 95 percent confidence bands.

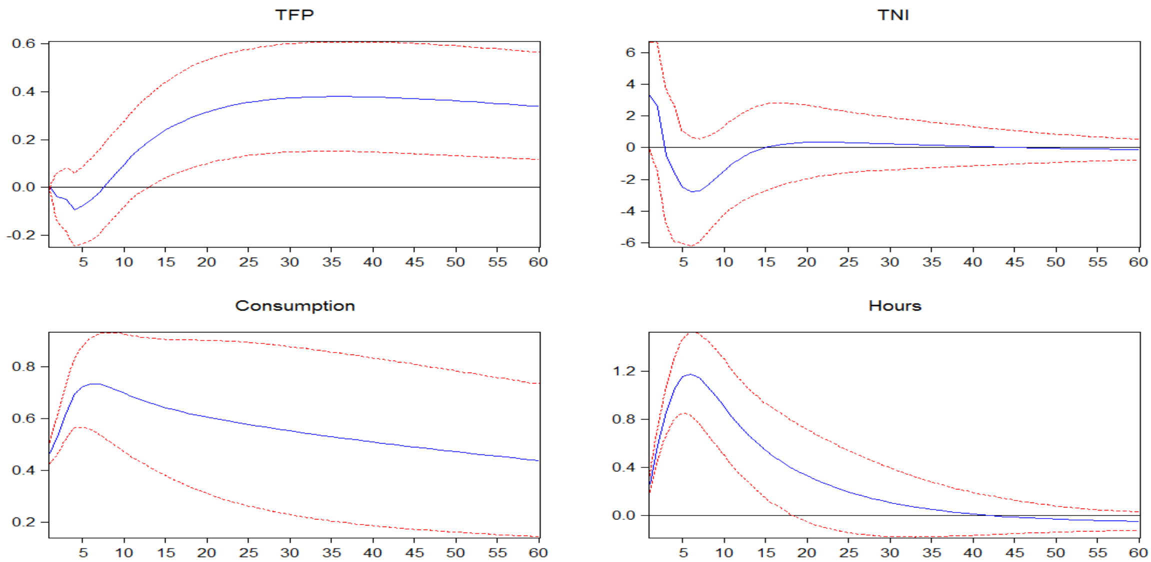

Figure 19: Impulse response to a news shock, ϵ_2, in the benchmark SVAR estimation with an exogenous dummy variable.

Notes: one-standard deviation ϵ_2 shock in percent deviation from the steady state. The blue solid line represents the point estimate and red dashed lines represent 95 percent confidence bands.

6 Conclusions

This paper presents a novel measure of news, derived from the relative frequency in which articles about technology and new developments are published in a major US news outlet, and use it to identify technology news shocks. Constructed technology news index is a suitable variable for the identification of news shocks, as it carries information about future productivity developments, and is able to produce a S-shaped response of TFP following the shock, consistent with slow diffusion of innovations view.

Furthermore, innovations to the index orthogonal to TFP lead to a boom in the economy and strong comovements in the selected macroeconomic variables immediately after technology news arrive. This takes place preceding actual productivity gains, consistent the news-driven business cycles view. Moreover, the shock identified with short and long-run restrictions follows very closely a shock that drives long-run paths of productivity.

The paper shows that changes in expectation due to arrival of news about future productivity are major driving forces of macroeconomic fluctuations, accounting for more than half of output's forecast error variance at business cycles frequencies. Additionally, the paper finds that newspaper articles about technology and science carry information about technological developments and, ultimately, about future productivity. Such results is one step towards research on how mass media coverage of tech and science events can influence agents' expectations.

References

Alexopoulos, M. and Cohen, J. (2015). The power of print: Uncertainty shocks, markets, and the economy. *International Review of Economics & Finance*, 40(C):8–28.

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

Barberá, P., Boydstun, A. E., Linn, S. F., McMahon, R. B., and Nagler, J. (2016). Methodological challenges in estimating tone: Application to news coverage of the U.S. economy.

Barro, R. J. and King, R. G. (1984). Time-Separable Preferences and Intertemporal-Substitution Models of Business Cycles. *The Quarterly Journal of Economics*, 99(4):817–839.

Barsky, R. B., Basu, S., and Lee, K. (2014). Whither News Shocks? In *NBER Macroeconomics Annual 2014, Volume 29*, NBER Chapters, pages 225–264. National Bureau of Economic Research, Inc.

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

Barsky, R. B. and Sims, E. R. (2012). Information, Animal Spirits, and the Meaning of Innovations in Consumer Confidence. *American Economic Review*, 102(4):1343–1377.

Beaudry, P. and Lucke, B. (2010). Letting Different Views about Business Cycles Compete. In *NBER Macroeconomics Annual 2009, Volume 24*, NBER Chapters, pages 413–455. National Bureau of Economic Research, Inc.

Beaudry, P. and Portier, F. (2004). An exploration into Pigou's theory of cycles. *Journal of Monetary Economics*, 51(6):1183–1216.

Beaudry, P. and Portier, F. (2005). The news view of economic fluctuations: Evidence from aggregate Japanese data and sectoral US data. *Journal of the Japanese and International Economies*, 19(4):635–652.

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

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

Cascaldi-Garcia, D. (2019). *Forecast revisions as instruments for news shocks*. Unpublished manuscript, Federal Reserve Board, Washington, DC. Available at `https://sites.google.com/site/cascaldigarcia/research`.

Christiansen, L. E. (2008). Do Technology Shocks Lead to Productivity Slowdowns? Evidence from Patent Data. IMF Working Papers 08/24, International Monetary Fund.

Cochrane, J. H. (1994). Shocks. *Carnegie-Rochester Conference Series on Public Policy*, 41(1):295–364.

Damstra, A., Boukes, M., and Vliegenthart, R. (2018). The economy. How do the media cover it and what are the effects? A literature review. *Sociology Compass*, 12(5):e12579.

Faccini, R. and Melosi, L. (2018). The Role of News about TFP in U.S. Recessions and Booms. Working Paper Series WP-2018-6, Federal Reserve Bank of Chicago.

Fare, R., Grosskopf, S., Norris, M., and Zhang, Z. (1994). Productivity Growth, Technical Progress, and Efficiency Change in Industrialized Countries. *American Economic Review*, 84(1):66–83.

Fernald, J. G. (2014). *A quarterly, utilization-adjusted series on total factor productivity.* Federal Reserve Bank of San Francisco Working Paper 2012-19. Technical report.

Forni, M., Gambetti, L., Lippi, M., and Sala, L. (2017). Noisy News in Business Cycles. *American Economic Journal: Macroeconomics*, 9(4):122–152.

Forni, M., Gambetti, L., and Sala, L. (2014). No News in Business Cycles. *Economic Journal*, 124(581):1168–1191.

Gentzkow, M., Kelly, B. T., and Taddy, M. (2017). Text as Data. NBER Working Papers 23276, National Bureau of Economic Research, Inc.

Gordon, R. J. (2013). U.S. Productivity Growth: The Slowdown Has Returned After a Temporary Revival. *International Productivity Monitor*, 25:13–19.

Gort, M. and Klepper, S. (1982). Time Paths in the Diffusion of Product Innovations. *Economic Journal*, 92(367):630–653.

Grimmer, J. and Stewart, B. M. (2013). Text as Data: The Promise and Pitfalls of Automatic Content Analysis Methods for Political Texts. *Political Analysis*, 21(03):267–297.

Jaimovich, N. and Rebelo, S. (2009). Can News about the Future Drive the Business Cycle? *American Economic Review*, 99(4):1097–1118.

Jorgenson, D. W., Ho, M. S., and Samuels, J. D. (2014). Long-term estimates of us productivity and growth. In *Third World KLEMS Conference. Growth and Stagnation in the World Economy, Tokyo.*

Keynes, J. M. (1936). *The General Theory of Employment, Interest and Money.* Macmillan. 14th edition, 1973.

Kurmann, A. and Mertens, E. (2014). Stock Prices, News, and Economic Fluctuations: Comment. *American Economic Review*, 104(4):1439–1445.

Larsen, V. H. and Thorsrud, L. A. (2019a). Business Cycle Narratives. CESifo Working Paper Series 7468, CESifo Group Munich.

Larsen, V. H. and Thorsrud, L. A. (2019b). The value of news for economic developments. *Journal of Econometrics*, 210(1):203 – 218.

Miranda-Agrippino, S., Hoke, S. H., and Bluwstein, K. (2018). When Creativity Strikes: News Shocks and Business Cycle Fluctuations. Discussion Papers 1823, Centre for Macroeconomics (CFM).

Nordhaus, W. (2004). Retrospective on the 1970s Productivity Slowdown. NBER Working Papers 10950, National Bureau of Economic Research, Inc.

Pigou, A. (1927). *Industrial Fluctuations.* London: Macmillan and Co.

Portier, F. (2014). Comment on "Whither News Shocks?". In *NBER Macroeconomics Annual 2014, Volume 29*, NBER Chapters, pages 265–278. National Bureau of Economic Research, Inc.

Ramey, V. A. (2016). Macroeconomic Shocks and Their Propagation. NBER Working Papers 21978, National Bureau of Economic Research, Inc.

Rogers, E. M. (2003). *Diffusion of innovations.* Free Press, New York, NY [u.a.], 5th edition.

Rotemberg, J. J. (2003). Stochastic Technical Progress, Smooth Trends, and Nearly Distinct Business Cycles. *American Economic Review*, 93(5):1543–1559.

Saint-Paul, G. (1993). Productivity growth and the structure of the business cycle. *European Economic Review*, 37(4):861–883.

Schmitt-Grohé, S. and Uribe, M. (2012). What's News in Business Cycles. *Econometrica*, 80(6):2733–2764.

Shiller, R. J. (2017). Narrative Economics. *American Economic Review*, 107(4):967–1004.

Smets, F. and Wouters, R. (2007). Shocks and Frictions in US Business Cycles: A Bayesian DSGE Approach. *American Economic Review*, 97(3):586–606.

Soroka, S. N., Stecula, D. A., and Wlezien, C. (2015). It's (Change in) the (Future) Economy, Stupid: Economic Indicators, the Media, and Public Opinion. *American Journal of Political Science*, 59(2):457–474.

Stock, J. H. and Watson, M. W. (2018). Identification and Estimation of Dynamic Causal Effects in Macroeconomics Using External Instruments. *The Economic Journal*, 128(610):917–948.

Tetlock, P. C. (2007). Giving Content to Investor Sentiment: The Role of Media in the Stock Market. *Journal of Finance*, 62(3):1139–1168.

Appendices

A List of subject keywords

The complete list of keywords used in the selection comprises 228 keywords and it is presented below:

- 1st ussr earth satellite - details on flight data obtained - other scientific data
- 2d ussr earth satellite - details on flight data obtained - other scientific data
- 2d ussr earth satellite - details on flight dog passenger data obtained - other scientific data
- 2d ussr earth satellite - details on launching dog passenger equipment tracking data transmission - other scientific data
- 2d ussr earth satellite: details on flight dog passenger data obtained - other scientific data
- academic and scientific journals
- adaptive technology
- advanced communications technology satellite
- advanced research projects agency
- advanced technologies
- aeronautical sciences journal of
- aerospace industries and sciences
- aerospaceindustries and sciences
- air technology international congress on
- albany medical center prize in medicine and biomedical research
- arms research and development
- arms research development and tests
- arthritis care and research (journal)
- arts and sciences inst
- association for competitive technology
- atomic energy and nuclear research
- ats (applications tech satellite) 1
- ats (applications technology satellite) 1
- ats (applications technology satellite) 2
- ats (applications technology satellite) 3
- ats (applications technology satellite) 4
- ats (applications technology satellite) 5
- ats (applications technology satellite) project
- ats (applications technology satellites) project
- aviation week & space technology (pub)
- awards for research and service
- awards for research and services
- bartlett tree research laboratories
- belgrade internatl fair of technics
- biotechnology
- biotechnology and bioengineering
- blockchain (technology)
- bluetooth (broadcasting technology)
- bluetooth wireless technology
- bulletin of the atomic scientists
- carnegie corp program for research insts

- chemical warfare by us opposed by 22 us scientists
- coal research bituminous
- communications/ research/ machines
- communications/ research/machines
- d technology
- earth science
- east river science park (nyc)
- edinburgh science festival
- electro-technology (pub)
- ence and technology
- encyclopedia of sports sciences and medicine
- environmental science services administration
- eur orgn for nuclear research (geneva)
- federal science progress (pub)
- forensic science
- forensic sciences
- foxconn technology
- general research and speculation
- ger scientists use by allies
- glendale technology park (endicott ny)
- govt research programs (hr select com) (elliott)
- gruss magnetic resonance research center (nyc)
- hall of science (flushing queens)
- health research and educational trust
- higgins scientific trust
- imaging (science specialty)
- imaging science
- indus research
- industrial research
- industrial research (pub)
- industrial research conference on
- industrial research laboratories
- information technologies building (newark nj)
- intel international science and engineering fair
- intel science talent search
- international science and engineering fair
- international science and technology (pub)
- inventions
- inventions and inventors
- inventions and patents
- inventions andinventors
- knorr (research vessel)
- laboratories and scientific equipment
- lasker albert medical research award
- li (westbury) fair and science and indus expositions
- life sciences
- materials testing and research
- med training and research
- medal of science (us)
- medal of technology (us)
- medical research planning conference
- micron technology inc
- military art and science
- millinery research (pub)
- nanotechnology
- national goals research staff
- national medal of science
- naval research special devices seminar
- neurosciences
- neursciences institute (la jolla calif)
- new models
- new models and design
- new models and designs
- new models and supply
- new models design & products
- new models design and product
- new models design and products

- new models design andproducts
- new models designand products
- new modelsdesign and products
- new technology telescope
- new us patents
- new york hall of science playground (queens)
- night vision technology
- nuclear research
- nuclear research ed on device developed by hofstadter and assocs
- nynex science and technology award
- other research and speculation
- other research speculation and exhibits
- patents
- patents (92%)
- patents granted
- pew research center poll
- popular science monthly (pub)
- process research products
- progress center research park (alachua fla)
- reasearch and technical developments
- rensselaer technology park (troy ny)
- research
- research (gen)
- research (general)
- research (sp
- research and development
- research and developments
- research and new developments
- research and patents
- research and tech development
- research and tech developments
- research and technical development
- research and technical developments
- research and therapy
- research development and tests
- research development and tests (gen)
- research funds for
- research grants
- research grants fed
- research tech developments
- research technical developments and theory
- research uses
- research uses of animals
- research:
- researchdevelopment and tests
- robinson technical products
- rocket science and propulsion
- rocky mountain research station
- science
- science (pub)
- science 80 (pub)
- science a
- science and
- science and life (pub)
- science and scientists
- science and technology
- science and technology (journal)
- science and technology (pub)
- science and technology science and technology
- science and world affairs conf on (coswa)
- science andtechnology
- science conference alaskan
- science congress pacific
- science digest (pub)
- science engineering and technology panel intergovernmental
- science illustrated (pub)
- science information exchange
- science news (pub)
- science service
- science talent search

- science week (pub)
- scienceand technology
- scientific american (pub)
- scientific freeom and responsibility award
- scientific information international conference on
- scientific research and development inter-dept com on
- scientists conf internatl (gb)
- scientists in govt
- scientists role
- serendip (scientific project)
- shamsher prakash research award
- siemens westinghouse math science and technology competition
- siemens westinghouse science and technology competition
- space technology laboratories
- speculations in science and technology (pub)
- standard & poor's high technology composite index
- tech aid conf
- tech developments and research
- technetium
- technetium (tc)
- technical data control
- technion (pub)
- technology
- technology entertainment and design conference
- theory and pure research
- therapy and research
- third international mathematics and science study
- training and supply of scientist engineers and allied technologists
- training and supply of scientists engineers and allied technologists
- training and supply of scientists engineers and technologists
- uber technologies inc
- united technologies corporation
- university heights science park (newark nj)
- upper atmosphere research satellite
- us and ussr satellite plans - satellite launched by ussr - details on launching tracking data transmission - other scientific data - speculation on future space flights
- us atomic scientists confs with officials
- us natl science foundation (proposed)
- us new models
- us patents
- us patents granted
- us research
- us research and development
- ussr earth satellite - details on flight tracking data transmission - other scientific data
- us-ussr joint research
- van nostrand's scientific encyclopedia
- westinghouse science awards
- westinghouse science talent search
- westinghousescience talent search
- world science festival
- world science festival (nyc)

B Additional results

B.1 Sentiment-based index

In this appendix, we intend to measure the sentiment contained in the technology-related articles that are published in the The New York Times. The calculation of the sentiment index is straightforward and follows the approach used by Tetlock (2007) and Soroka et al. (2015) for capturing the sentiment contained in newspaper corpus: after selecting the articles of interest based on the list of keywords, the sentiment associated with each article is calculated from its correspondent lead paragraph as the number of positive words minus the number of negative keywords divided by the total number of words contained in the article. Finally, the sentiment associated with each quarter of the sample is calculated by taking the mean of the sentiment scores of all the articles published in that given quarter. The bag of positive and negative words is retrieved from the *Harvard IV-4 Dictionary*, a general-purpose dictionary developed by the Harvard University. For a given quarter t from the sample, with N technology-related articles contained in it, the Sentiment Score is calculated as shown in Equation 24 and the raw Technology Sentiment Index (TSI) is depicted in Figure 20.

$$S_t = \frac{\sum_{i=1}^{N_t}\left(\frac{(\#positivewords_i - \#negativewords_i)}{\#totalwords_i}\right)}{N_t}, \tag{24}$$

Next, as a way to compare the utility of TSI for the identification of news shocks, the 4-variable specification of the benchmark model from section 5 is estimated with the sentiment-based index in place of the frequency-based one, the TNI.

Figure 21 depicts the impulse responses of TFP, TNI, consumption (C) and hours worked (H) to a shock $\hat{\epsilon}_2$ that was identified as the innovation to TSI that is orthogonal to TFP, using the same identification scheme from the benchmark model with TNI. The impulse response functions show that the shock $\hat{\epsilon}_2$ identified is not in accordance with what is expected from a news shock: TFP does not have a large response in the short-run, however, it stars respond negatively after 9 quarters; moreover, the responses of consumption and hours worked does match with the predictions of news-driven business cycles advocates nor to the traditional RBC models, i.e., both variables decrease after a news shock, and remain negative for a long period.

Therefore, sentiment-based indexes can, to some extent, capture the "animal spirits" of the agents. For this paper, though, it did not show to be the most appropriate approach for proper identification of news shocks.

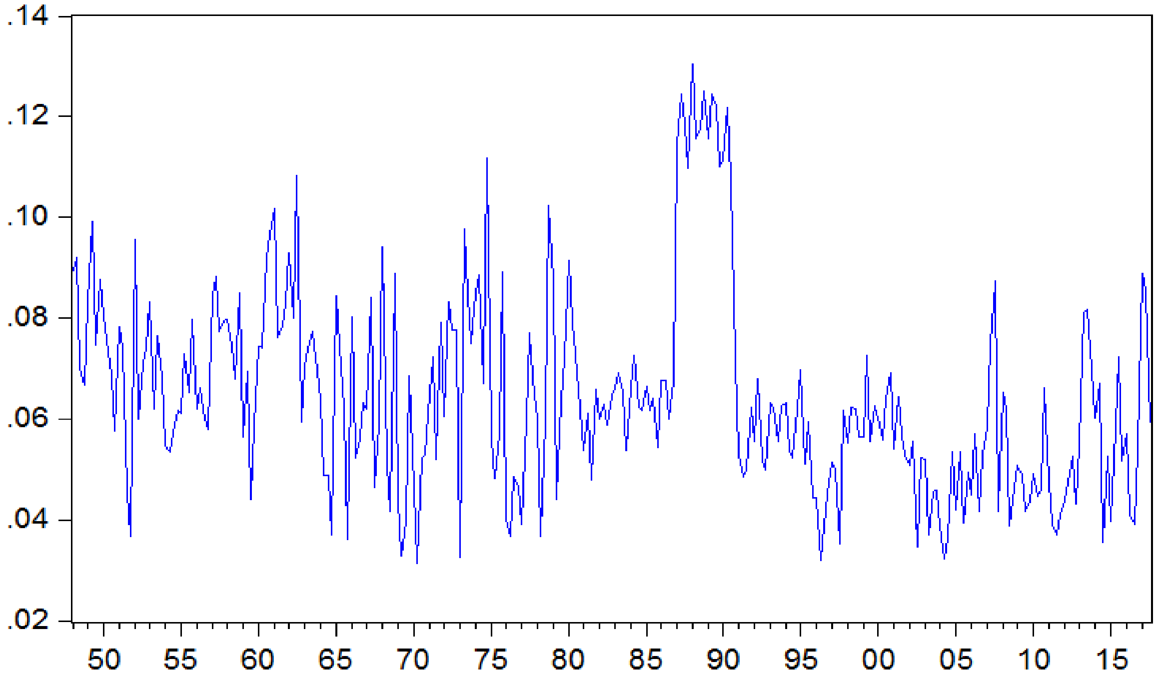

Figure 20: Technology Sentiment Index (TSI).

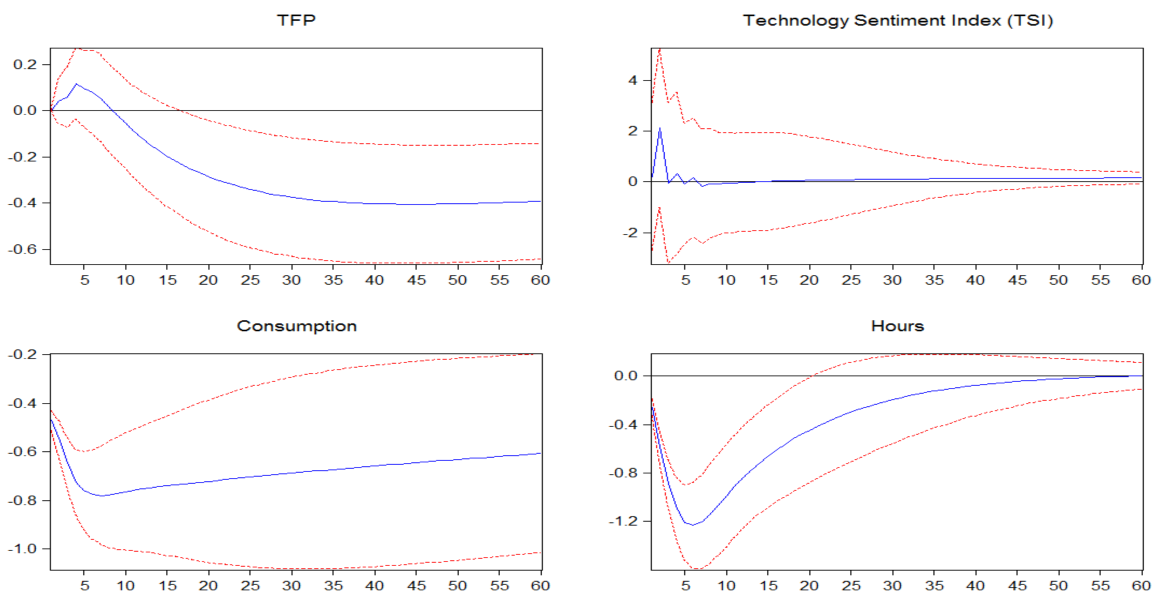

Figure 21: Impulse responses to a sentiment shock, $\hat{\epsilon}_2$, in the 4-variable system $[TFP_t\ TSI_t\ C_t\ H_t]$.

Notes: Each graph in the figure depicts the impulse response to a one-standard deviation $\hat{\epsilon}_2$ shock in percent deviation from the steady state. The blue line represents the point estimate of the Impulse Response Functions. The red dotted lines represent 95-percent confidence bands.

Non-exclusive licence to reproduce thesis and make thesis public

I, **Caio Augusto Raduy Koslyk**,

1. herewith grant the University of Tartu a free permit (non-exclusive licence) to reproduce, for the purpose of preservation, including for adding to the DSpace digital archives until the expiry of the term of copyright,

News-Driven Business Cycles: A Narrative Approach,

supervised by **Lenno Uusküla**.

2. I grant the University of Tartu a permit to make the work specified in p. 1 available to the public via the web environment of the University of Tartu, including via the DSpace digital archives, under the Creative Commons licence CC BY NC ND 3.0, which allows, by giving appropriate credit to the author, to reproduce, distribute the work and communicate it to the public, and prohibits the creation of derivative works and any commercial use of the work until the expiry of the term of copyright.

3. I am aware of the fact that the author retains the rights specified in p. 1 and 2.

4. I certify that granting the non-exclusive licence does not infringe other persons' intellectual property rights or rights arising from the personal data protection legislation.

Caio Augusto Raduy Koslyk

Tartu, 23.05.2019