# The Network Effects of Fiscal Adjustments*

Edoardo Briganti† Carlo A. Favero‡
Madina Karamysheva§

June 10, 2021

**Abstract**

*We study the effects of fiscal consolidations in the United States and their propagation in the production network. We use a narrative approach to identify fiscal adjustments which are exogenous to output fluctuations. Then we apply spatial econometric techniques to separate the total effect of fiscal adjustments into a direct and network component. We find that fiscal adjustments based on increased taxation are more recessionary than those based on spending cuts. Moreover, one quarter of the difference in their total output effect is explained by the stronger network propagation of taxes relative to government spending.*

**Keywords:** industrial networks, fiscal adjustment plans, output growth, applied spatial econometrics.
**JEL codes :** E60, E62.

---

*This paper was presented during IAAE 2018 conference meeting, HSE and Bocconi seminars. We are grateful to participants for useful comments; in particular, we thank Roberto Perotti, Hashem Pesaran, Lung-Fei Lee and Valerie Ramey for useful comments.

†PhD student at University of California San Diego, ebrigant@ucsd.edu

‡Deutsche Bank Chair, IGIER-Bocconi, and CEPR (full address: IGIER and Dept. of Finance, Universit Commerciale Luigi Bocconi, via Roentgen 1, 20136 Milano Italy), tel: +390258363306; carlo.favero@unibocconi.it

§NRU Higher School of Economics, Moscow (full address: Faculty of Economic Sciences, National Research University Higher School of Economics, Pokrovsky Boulevard 11, office number S-639, Moscow, Russian Federation), tel: +79192370529; mkaramysheva@hse.ru

# Contents

# 1 Introduction

In a world overwhelmed by a global health pandemic, the adoption of unprecedented fiscal stimuli has deteriorated the state of public finances. Economic theory and good practice suggest that a government should run a deficit during recessions, when tax revenues are low and government spending is high due to fiscal stabilizers like unemployment subsidies. The same holds during periods of temporarily high spending needs, when a government must cope with catastrophes such as financial crises, natural disasters, or wars. These deficits should be balanced by surpluses during economic booms and when spending needs are low. As global economies recover from the COVID-19 crisis and return to growth, fiscal consolidations play a crucial role in balancing government budgets and bringing sovereign debt below the "maximum sustainable debt" threshold (see e.g., Collard, Habib, and Rochet (2015)).[1] In this context, understanding the transmission mechanism and output effects of fiscal consolidations is crucial for policymakers hoping to design optimal fiscal adjustment plans.[2]

As recently reported in Ramey (2019), the fiscal policy literature has consistently found a new empirical fact: fiscal consolidations due to higher taxes imply larger output losses compared to consolidations due to reductions in government spending. This pattern has been confirmed in recent studies which compare the effects of austerity measures across a panel of countries (see e.g., Guajardo, Leigh, and Pescatori (2014) and Alesina, Favero, and Giavazzi (2015)). However, there is little work investigating the underlying reasons for such an asymmetric response.

In this paper, we explore production networks as a potential explanation for the asymmetric output response of fiscal consolidations. Contrary to existing literature, we restrict our analysis to one single country, namely, the United States. This has two main benefits. Firstly, we are able to estimate effects specific to the US and thus provide more reliable guidance from a policy-making perspective, as multi-country analysis tends to report country-average effects. Secondly, we are able to exploit rich industry-level data to track of the effects of fiscal consolidations at a more disaggregated level. As a result, we shed light on the transmission mechanism of fiscal policy and quantify its propagation through the industrial network.

[1] On April 25th 2020, in an article entitled "After the disease, the debt", The Economist wrote: "... governments should prepare for the grim business of balancing budgets later in the decade."

[2] For a literature survey on sovereign debt, see Panizza, Sturzenegger, and Zettelmeyer (2009) and Reinhart and Rogoff (2009)

In particular, we motivate our work with the following questions. What are the effects of fiscal adjustments in the United States? Are tax-based fiscal consolidations more recessionary than expenditure-based consolidations, as highlighted by the country-panel analysis? Can asymmetries in the input-output network explain this difference? To provide an answer, we study the effects of fiscal consolidations implemented in the Unites States from 1978 to 2014. These effects propagate through a 62-industry production network.
Regarding the first two questions, we find that tax-based (TB henceforth) fiscal adjustments have a recessionary output multiplier over two years of -1.27% while the effects of expenditure-based (EB henceforth) fiscal plans are not statistically different from zero. These results are in line with those obtained by the current state of the literature which uses a panel of OECD countries.
Moreover, we answer the third question using cutting-edge spatial econometric techniques, assuming that the observed units, 62 industries, are "spatially connected" via an input-output production network. The spatial framework allows us to decompose the aggregate total effects of fiscal consolidations into a direct and a network effect. The former represents the direct impact of the fiscal shock on each industry while the latter represents the spillover effects from other industries hit by the same aggregate shock. In turn, we are able to investigate if the stronger recessionary effects of TB fiscal consolidations relative to EB are explained by differences in the network propagation mechanism of these shocks.

Our baseline results suggest that 27% of the total effect of TB fiscal consolidations come from network spillovers. This result is robust to various specifications. On the other hand, network effects of EB plans are more modest and less robust, with only 11% of the total output effect coming from the network. Overall, the stronger network effect of TB plans explains close to one-fourth of the differences in the total effects of TB and EB plans. Networks thus provide a partial explanation of asymmetry in the output response of these two types of fiscal consolidations.
In addition to these results, this paper has two other original contributions. To the best of our knowledge, we are the first to study and detect input-output spillovers of taxes. We find that network propagation of tax shocks is driven by a few key suppliers in the economy.[3] This result is consistent with Ozdagli and Weber (2017), who study upstream propagation of monetary policy shocks. As noted in Ozdagli and Weber (2017), spatial models of the macro-economy are

[3] Key suppliers in the network are: Fabricated Metal Products, Primary Metals, Wholesale Trade, Plastic and Rubber Products, Chemical Products, Real Estate, Administrative Services, Miscellaneous Professional, Scientific and Technical Services.

a useful tool for understanding the sources and transmission mechanism of aggregate shocks. As far as we know, this paper and Ozdagli and Weber (2017) are the only two studies which make use of this class of models. In addition, we provide some practical guidance on using spatial econometric techniques to answer macroeconomic questions.

### Related Literature

First of all, our paper relates to the literature of fiscal consolidations: Guajardo, Leigh, and Pescatori (2014), Alesina, Favero, and Giavazzi (2015) and Alesina, Barbiero, et al. (2017). Unlike these papers, we consider a panel of US industries rather than countries, and we are the first to study the network effects of fiscal consolidations.

Alesina, Barbiero, et al. (2017) also propose a theoretical explanation for the stronger effect of TB fiscal consolidations. On the other hand, we provide an alternative, network-based explanation of the asymmetric output effects of TB and EB plans.

Secondly, our work relates to the seminal works of Gabaix (2011), Acemoglu, Carvalho, et al. (2012), which develop the role of production networks in amplifying the effects of localized shocks.[4] In Section 3.2, we connect our empirical findings to the network literature. However, unlike these papers, our work adopts a spatial framework to determine the extent to which the total effects of fiscal policy can be attributed to network transmission. This point has been highlighted in Ozdagli and Weber (2017), who perform a similar analysis to study the propagation of monetary policy shocks in the US stock market.

Acemoglu, Akcigit, and Kerr (2016) study the asymmetric propagation in the production network of demand and supply shocks. In particular, they find that government spending shocks uniquely propagate upstream in the production network, from customers to suppliers. Bouakez, Rachedi, Emiliano, et al. (2018) find that government spending has larger effects when it is concentrated in sectors located downstream in the production network, which account for a smaller share of private final demand. Unlike them, we study empirically the propagation of a special type of fiscal shock, namely TB and EB fiscal consolidations.

Thirdly, this work relates to the literature on fiscal policy at an industry level:

[4] Other recent theoretical and empirical contributions include, but not limited to Baqaee and Farhi (2018), Baqaee and Farhi (2019a), Baqaee and Farhi (2019b), Barrot and Sauvagnat (2016), Boehm, Flaaen, and Pandalai-Nayar (2019). Carvalho and Tahbaz-Salehi (2019) summarize the literature providing both theoretical foundation for production networks as a propagation channel as well as evidence from growing empirical literature.

Ramey and Shapiro (1999), Perotti (2007) and Nekarda and Ramey (2011). In particular, Nekarda and Ramey (2011) focus on government purchases in manufacturing industries and find evidence in support of the Neo-Classical model. They also construct a comprehensive measure of government purchases which takes into account downstream linkages. Building on this work, we provide analysis of the transmission mechanism of fiscal policy at an industry level. Additionally, we enrich this analysis by using all the industries in the economy and by integrating them into a production network.
Cox et al. (2020) study public procurement contracts and find large sectoral bias in government spending. Our industry analysis thus takes into account the sectoral heterogeneity of fiscal policy effects. Auerbach, Gorodnichenko, and Murphy (2019) use city-level data on local defense public procurement and find large fiscal (first-order) spillovers among industries. Their results contradict our finding of weak propagation of EB plans. However, it is hard to provide a direct comparison between our two results since we use different levels of aggregation and defense public procurement contracts are different from EB fiscal adjustment plans.

The rest of the paper is organized as follows. In Section 2, we illustrate how fiscal adjustment plans identify exogenous fiscal consolidation policies. This section also introduces our spatial auto-regression equation and our data. Section 3 summarizes results and rationalizes them with economic theory, and then concludes with some brief guidelines for macroeconomists interested in using spatial econometrics for their research. Section 4 provides some robustness checks and Section 5 concludes.

# 2 Empirical Approach and Data

## 2.1 Identification of Exogenous Fiscal Adjustments Plans

Measuring the propagation of fiscal adjustments requires the identification of an exogenous demand and supply shocks. Our identification strategy thus relies on the narrative analysis of fiscal adjustment plans. This strategy is a recent innovation in the fiscal policy literature and employs narrative exogenous shocks as a proxy for fiscal consolidation policies. This strategy was introduced in Alesina, Favero, and Giavazzi (2015) to take into account the fact that fiscal adjustments are implemented through multi-year plans with both an intertemporal and an intratemporal dimension.
The intratemporal dimension refers to the fact that fiscal consolidations are

implemented with a mix of tax increases and spending cuts. Tax and the expenditure components of the adjustments are correlated since governments decide first on the size of the adjustment, and then on its composition in terms of expenditures and revenues. The intertemporal dimension is important since fiscal adjustments are implemented via multi-year plans with measures upon announcement (the unanticipated component of the plan) and measures announced for subsequent years (the anticipated component of the plan). In particular, each country has a specific "recipe" to implement fiscal consolidations: some countries prefer to unexpectedly raise taxes without cutting expenditures, while others announce large future cuts in spending and only marginally increase taxes. Alesina, Favero, and Giavazzi (2015) refer to this as the country-specific "style of the plan".

These complications make identifying pure and isolated tax hikes and spending cuts during years of fiscal consolidation a difficult, if not impossible, task. Fiscal plans provide an effective tool to circumvent these difficulties when studying austerity policies.

**Modeling Fiscal Plans:**

From a mathematical standpoint, plans are sequences of fiscal corrections, announced at time $t$ and implemented between $t$ and $t+K$, where $K$ is the anticipation horizon. In each year $t$, two types of fiscal corrections are possible:

1. The *unanticipated fiscal shock*, that is, the surprise change in the primary surplus at time $t$, which we denote by:

$$f_t^u := tax_t^u + exp_t^u,$$

   where $tax_t^u$ is the surprise increase in taxes announced and implemented at time $t$, while $exp_t^u$ is the surprise reduction in government expenditure also announced and implemented at time $t$.

2. The *anticipated fiscal shock*: the change in the primary surplus at time $t$, which had already been announced in the previous years and is either implemented in year $t$ or scheduled to happen within $K$ years. In particular, we denote as $tax_{t,j}^a$ and $exp_{t,j}^a$ the tax and expenditure changes announced by the fiscal authorities at date $t$ with an anticipation horizon of $j$ years (*i.e.,* to be implemented in year $t+j$). Therefore, we further distinguish between:

   (a) The *anticipated implemented shock*: scheduled in the past and implemented in year $t$:

$$f_t^a := \; tax_{t,0}^a + exp_{t,0}^a$$

(b) The <u>*anticipated future shocks*</u>: sum of scheduled tax and government spending changes which have to be implemented within $K$ years from their announcement:

$$f_t^f := \sum_{j=1}^{K} tax_{t,j}^a + \sum_{j=1}^{K} exp_{t,j}^a.$$

In a fiscal adjustment database, as long as no policy revision takes place, the anticipated shocks roll over year-by-year. In formulae:

$$tax_{t,j}^a = \underbrace{tax_{t-1,j+1}^a}_{\textit{Old shock, rolled over}} \qquad exp_{t,j}^a = \underbrace{exp_{t-1,j+1}^a}_{\textit{Old shock, rolled over}} .$$

However, if from one year to another, a policy revision takes place, then, the new anticipated future shock will embed such change:[5]

$$\begin{aligned} tax_{t,j}^a &= \underbrace{tax_{t-1,j+1}^a}_{\textit{Old shock, rolled over}} + \underbrace{\left(tax_{t,j}^a - tax_{t-1,j+1}^a\right)}_{\textit{Policy Revision}}, \quad \text{with } j \geq 1 \\ exp_{t,j}^a &= \underbrace{exp_{t-1,j+1}^a}_{\textit{Old shock, rolled over}} + \underbrace{\left(exp_{t,j}^a - exp_{t-1,j+1}^a\right)}_{\textit{Policy Revision}}, \quad \text{with } j \geq 1 \end{aligned}$$

We adopt the annual database on fiscal adjustment plans constructed by Alesina, Favero, and Giavazzi (2015) and consider only fiscal consolidations that occurred in the US from 1978 to 2014. They identify fiscal adjustments exogenous with respect to output fluctuations using a narrative identification method. This approach is similar to C. D. Romer and D. H. Romer (2010), who identify exogenous tax shocks from presidential speeches, congressional debates, budget documents, and congressional reports. From these documents, they identify the size, timing, and principal motivation for all major postwar tax policy actions. Legislated changes are then classified into two categories: 1) endogenous, if induced by short-run counter-cyclical concerns; 2) exogenous, if taken in response to the state of government debt (deficit-driven). As mentioned earlier, fiscal adjustment plans allow us to control for the intertemporal and intratemporal correlation, which we report in Table I:

[5] In the above expression $j \geq 1$ since any policy revision introduced upon implementation ($j = 0$) is no longer a part of an anticipated shock; in fact, it is a new unanticipated component.

**Table I:** Inter- and Intra-temporal correlation matrix of Fiscal Adjustments Plans in the US

| | $tax_t^u$ | $tax_{t,0}^a$ | $tax_t^f$ | $exp_t^u$ | $exp_{t,0}^a$ | $exp_t^f$ |
|---|---|---|---|---|---|---|
| $tax_t^u$ | 1 | 0.041 | 0.570 | 0.596 | -0.126 | 0.105 |
| $tax_{t,0}^a$ | | 1 | 0.038 | 0.098 | 0.361 | 0.310 |
| $tax_t^f$ | | | 1 | -0.047 | 0.019 | 0.180 |
| $exp_t^u$ | | | | 1 | -0.050 | 0.014 |
| $exp_{t,0}^a$ | | | | | 1 | 0.782 |
| $exp_t^f$ | | | | | | 1 |

*Table I: linear correlation matrix of legislated changes in taxes and expenditure identified by the narrative analysis. Sample: annual data from 1978 to 2014 of US fiscal adjustment plans from Alesina, Favero, and Giavazzi (2015). In blue is reported the intra-temporal correlation (between each component of taxes and expenditures. In green is the inter-temporal correlation (within tax or expenditure component, but between components with different timing). In black we have a mix of the two: correlation between tax and expenditure components with different timing.*

Notice from Table I that the (intra-temporal) correlation between unanticipated tax and unanticipated expenditure adjustments is 60%. Similarly, the (inter-temporal) correlation between future and anticipated components of expenditure is 78%. As both the inter-temporal and the intra-temporal dimension matter, it is worth considering multi-year fiscal plans instead of individual measures of tax and government spending shocks.

Since this source of correlation confounds the effects of taxes and expenditures, we need to classify plans into mutually exclusive categories which can be simulated independently. We can then take into account the inter-temporal correlation within each category. To this end, we exploit the fact that not all the plans are the same. Some fiscal plans are designed to increase taxes more than cut expenditures and are labeled as TB (tax-based). On the contrary, those plans which rely more on expenditure cuts rather than tax hikes are labeled as EB (expenditure-based).

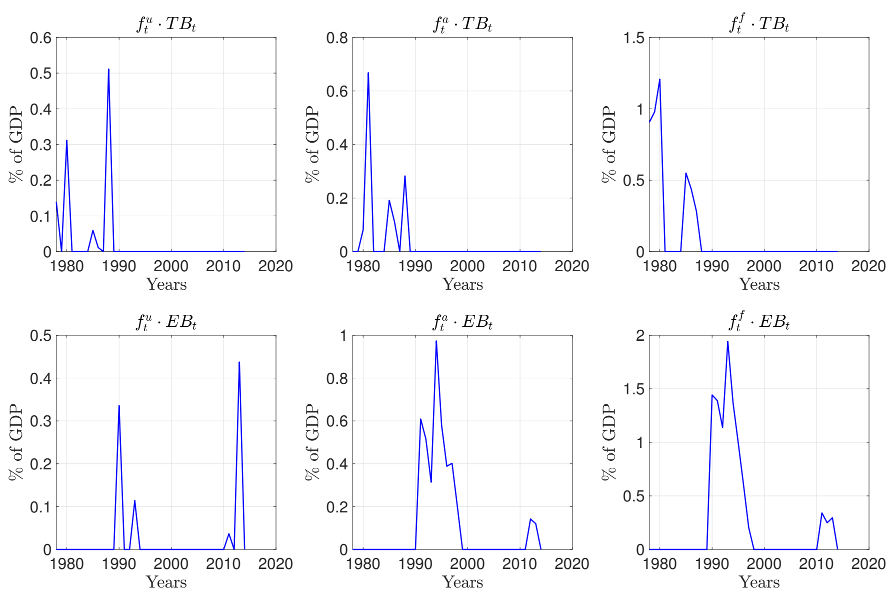
For instance, the criterion which determines whether a fiscal consolidation is labeled as TB can be written:

$$\underbrace{\left(tax_t^u + tax_{t,0}^a + \sum_{j=1}^{K} tax_{t,j}^a\right)}_{\text{overall tax hike in } t} > \underbrace{\left(exp_t^u + exp_{t,0}^a + \sum_{j=1}^{K} exp_{t,j}^a\right)}_{\text{overall expenditure cut in } t}. \tag{1}$$

Criterion (1) is saying that if the overall tax hike in year $t$ exceeds the overall spending cut, then we label year $t$ as a year of TB fiscal consolidation. We keep track of these years by constructing two dummy variables, $TB_t$ and $EB_t$, which are equal to one if year $t$ is labeled as TB or EB, respectively. By construction, TB and EB plans are mutually exclusive. That is, EB and TB plans cannot occur simultaneously. This lets us simulate separately the effect of TB and EB plans while preserving, within each type of plan, the observed intra-temporal correlation between adjustments on government's revenues and expenditure.

Figure 1 plots our fiscal adjustment database. This contains all of the nominal changes in taxes and expenditure, scaled by GDP of the year before the consolidation occurs to avoid potential endogeneity issues. Moreover, the future component of the fiscal adjustment plan has a maximum anticipation horizon of three years ($K$). This is in line with the small numbers of occurrences of policy shifts anticipated four and five years ahead, and is consistent with the database in Pescatori et al. (2011). The top row of Figure 1 illustrates the three components of fiscal adjustments interacted with the dummy $TB_t$ to identify the components of tax-based fiscal consolidations. The bottom row does the same for expenditure-based fiscal consolidations.

Figure 1: Fiscal Adjustments Plans - United States 1978-2014




We assess the goodness of our orthogonalization criterion (1), by showing in Figure 2 the share of tax increases and spending cuts of each total fiscal adjustment, $f_t^u + f_t^a + f_t^f$.

Figure 2: Fiscal Adjustment Composition

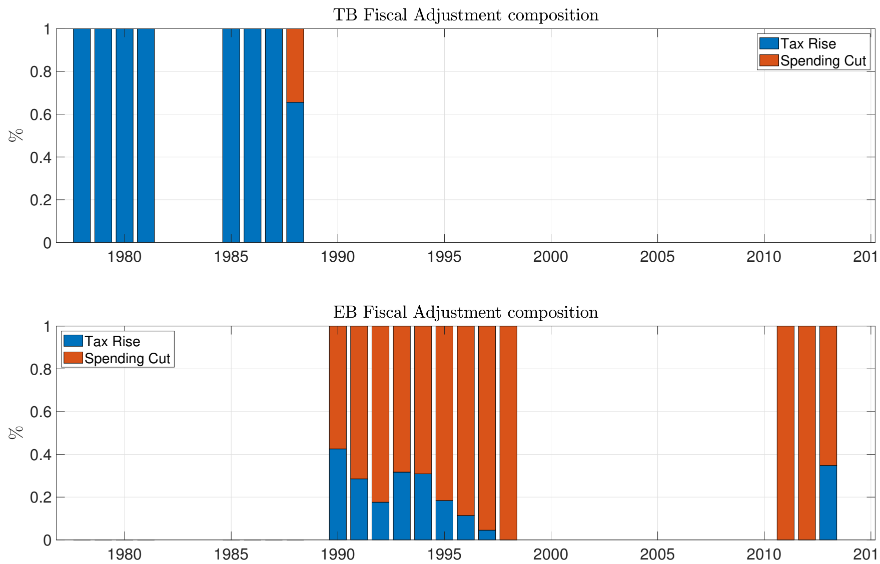



Figure 2 shows that the labeling of fiscal adjustment into EB or TB plans, by means of criterion (1), is never marginal: i. TB plans are all pure tax hikes except for the year 1988, which is the result of a hybrid fiscal plan with only 30% in spending cuts; ii. EB plans are mainly made up of spending cuts with only 20% of policy changes coming from a tax increase. Figure 2 also illustrates the timing of fiscal consolidations in the US: i. there are two periods of TB fiscal adjustments ($TB_t = 1$) between 1978-1981 and 1985-1988; ii. there are three periods of EB fiscal adjustments ($EB_t = 1$) between 1990-1992, 1993-1998 and 2011-2013.

To summarize, we classify fiscal consolidations into TB and EB fiscal adjustment plans to account for the observed correlation between tax and expenditure adjustments. This correlation comes from the fact that policy makers implement fiscal consolidations by adopting multi-year fiscal plans with both tax hikes and spending cuts.

Finally, we highlight that we are aware of the criticism that focusing only on tax increases and spending cuts might cause upward bias in the fiscal multipliers' estimates.[6] However, our goal is not to estimate standard fiscal multipliers. We are looking at the effects of fiscal consolidations, which are a mixture of tax hikes and spending cuts implemented for deficit reduction reasons. We

[6] We thank Valerie Ramey for making us aware of this potential pitfall.

are not censoring changes in "G" or "T". Rather, we are looking at a different variable, fiscal consolidations, denoted by $f_t$. Consequently, we do not claim that our estimates will be externally valid for a generic spending cut or tax increase. However, if the United States plans to undertake either a TB or an EB fiscal consolidation, our estimates are valid and can be used as a benchmark for policy-makers.

## 2.2 Regression Equation

Our econometric specification builds on Alesina, Favero, and Giavazzi (2015), who regress country-level output growth on the 3 components of TB and EB country-specific fiscal plans. Unlike them, we focus on a single country, the United States, by breaking down its economy into $n = 62$ industries. Furthermore, we enrich their specification with two *spatial variables* to take into account the input-output connections among sectors, and the spatial correlation generated by them. This is similar to the approach in Acemoglu, Akcigit, and Kerr (2016). Denoting in blue the parameters we estimate, we have:

$$\Delta \log y_{i,t} = \underbrace{\alpha_i}_{\textit{Industry FE}} + \rho^{down} \cdot \underbrace{\left( \sum_{j=1}^{n} \left( \frac{\text{SALES}_{j \to i}}{\text{SALES}_i} \right) \cdot \Delta \log y_{j,t} \right)}_{\Delta y_{i,t}^{down}} \cdot TB_t + \\ + \rho^{up} \cdot \underbrace{\left( \sum_{j=1}^{n} \left( \frac{\text{SALES}_{i \to j}}{\text{SALES}_i} \right) \cdot \Delta \log y_{j,t} \right)}_{\Delta y_{i,t}^{up}} \cdot EB_t + \boldsymbol{F}_{i,t}^T \cdot \boldsymbol{\beta} + \nu_{i,t} \qquad (2)$$

where $\boldsymbol{F}_{i,t}^T$ is the vector of fiscal adjustment plans identified in the previous section and $\nu_{i,t}$ is a serially uncorrelated, heteroskedastic error term. We allow for heteroskedasticity since sectors exhibit different volatility in growth rates in the data.

Given the existence of an intricate industrial network which links all the sectors, fiscal consolidations might easily spillover from one sector to another over a year time span. We refer to the spillover moving from customers to suppliers as upstream propagation, while the spillover from suppliers to customers is downstream propagation. In order to measure the effects of both transmission channels, we include in Equation (2) two spatial variables, namely $\Delta y_{i,t}^{\text{down}}$ and $\Delta y_{i,t}^{\text{up}}$. These variables are constructed as a weighted average of the fluctuations of customer industries (upstream propagation) and supplier industries (downstream propagation). The weights are given by the share of output sold

and purchased relative to an industry's output.
Intuitively, sectors are connected via input-output relationships which can propagate and amplify industry-specific shocks (see Acemoglu, Carvalho, et al. (2012)). For instance, suppose following an EB-type fiscal consolidation, the government decides to purchase less goods from sector X, a major customer of sector Y. Suppose also that Y does not procure goods to the government. In this scenario, the reduction of government demand for X does not directly impact Y. Yet, Y might still be indirectly harmed through X, if X shrinks its production and reduces the input purchased by Y. This is an example of upstream propagation, that is, the shock propagates from a directly affected customer sector upstream, thus propagating effects to a supplier industry. Similarly, shocks can propagate in the other direction. For instance, suppose sector X is heavily directly affected by a TB-type fiscal consolidation. Suppose also that sector X is a main supplier of input for sector Y. Even if Y might not be directly affected by the consolidation, if sector X decides to increase prices in response to the negative shock, this will indirectly hurt sector Y.
Equation (2) is implicitly assuming that TB fiscal consolidations exclusively propagate downstream, from suppliers to customers while the opposite is true for EB fiscal consolidations. This is done by interacting $TB_t$ with $\Delta y_{i,t}^{\mathrm{down}}$ and $EB_t$ with $\Delta y_{i,t}^{\mathrm{up}}$. This assumption is relaxed by switching the interaction and estimating also the following equation:

$$\Delta \log y_{i,t} = \tilde{\alpha}_i + \tilde{\rho}^{down} \cdot \Delta y_{i,t}^{down} \cdot EB_t + \tilde{\rho}^{up} \cdot \Delta y_{i,t}^{up} \cdot TB_t + \boldsymbol{F}_{i,t}^T \cdot \tilde{\boldsymbol{\beta}} + \tilde{\nu}_{i,t}. \quad (3)$$

By doing so we remain agnostic about the the direction of propagation of fiscal consolidations. Section 3.2 will provide a theoretical rationalization of our results.
Another option is to consider all propagation channels at once by estimating a single larger model which nests both Equation (2) and (3). This option is intractable due to the large number of parameters relative to the sample size, and due to collinearity between the spatial variables. We therefore estimate two separate models and then we apply a Vuong test for non-nested models to see which one fits the data better (see Vuong (1989) and Wooldridge (2010)).

Finally, we clarify the contents of the vector of fiscal consolidations:

$$\underset{1\times 6}{\boldsymbol{F}_{i,t}^{T}} \cdot \boldsymbol{\beta} = \begin{bmatrix} \begin{bmatrix} f_t^u & f_t^a & f_t^f \end{bmatrix} \cdot \omega_i^{TB} \cdot TB_t & \begin{bmatrix} f_t^u & f_t^a & f_t^f \end{bmatrix} \cdot \omega_i^{EB} \cdot EB_t \end{bmatrix} \cdot \begin{bmatrix} \tau_u \\ \tau_a \\ \tau^f \\ \gamma_u \\ \gamma_a \\ \gamma^f \end{bmatrix}$$

$$= \underbrace{\left(\tau_u \cdot f_t^u + \tau_a \cdot f_t^a + \tau^f \cdot f_t^f\right) \cdot \omega_i^{TB} \cdot TB_t}_{TB \textit{ Fiscal Consolidations}} + \underbrace{\left(\gamma_u \cdot f_t^u + \gamma_a \cdot f_t^a + \gamma^f \cdot f_t^f\right) \cdot \omega_i^{EB} \cdot EB_t}_{EB \textit{ Fiscal Consolidations}}$$

where $\omega_i^{TB}$ and $\omega_i^{EB}$ are industry specific weights, or "shifters" (using the shift-share design terminology), which distribute the aggregate shock among the different sectors. This is done to allow for heterogeneous effects of aggregate fiscal consolidations on each industry.

## 2.3 Industry Data

We focus on a partition of the US economy made by 62 industries, observed from 1978 to 2014 at a yearly frequency. The disaggregation level, 62, is determined by starting from the finest decomposition available on the Bureau of Economic Analysis (BEA) at a yearly frequency, namely 71 sectors, and then aggregating those sectors whose data are not available for older years.[7]

### Value Added

We use real industry value-added as the dependent variable, $\Delta y_{it}$. Value-added equals gross output minus intermediate inputs. It consists of compensation of employees, taxes on production and imports less subsidies (formerly indirect business taxes and non-tax payments), and gross operating surplus (formerly "other value added"). We prefer it over gross output to be consistent with Acemoglu, Akcigit, and Kerr (2016).[8]

[7] We exclude the Government sector and consider only Government Enterprises as the only public, but politically independent, sector. The Government sector needs to be excluded since its outcome variable is G, government spending, which mechanically falls when a fiscal adjustment occurs.

[8] Their decision is justified by the fact that value-added is adjusted for energy costs, non-manufacturing input, and inventory changes which are all outside of the general equilibrium model which provides the theoretical underpinning to their empirical strategy.

### Input-Output matrices

The BEA provides I-O tables that report the amount of commodity used (Use Table) and made (Make Table) by each industry. Horowitz, Planting, et al. (2006) outline the procedure to construct an industry-by-industry direct requirement matrix, with elements given by $SALES_{j \to i}/SALES_i$ for each sector.[9] Let's denote this matrix by $A$ and note that its elements coincide one to one with the weights of $\Delta y_{i,t}^{\text{down}}$ in Equation 2. Therefore, the downstream spatial variable can be written in vector notation as: $\Delta \boldsymbol{y}_t^{\text{down}} = A \cdot \Delta \log \boldsymbol{y}_t$ and matrix $A$ can be constructed from the Make and Use Tables of the BEA.[10] Henceforth we will refer to matrix $A$ as the "downstream matrix".

Finally, we construct a new matrix starting from $A$ and using BEA's industry specific gross output, such that its $(ij)_{th}$ element is represented by $SALES_{i \to j}/SALES_i$, which coincides one to one with the weights of $\Delta y_{i,t}^{\text{up}}$ in Equation 2. We denote this new matrix by $\hat{A}^T$, and refer to it as the "upstream matrix". The upstream spatial variable can now be written in vector notation as: $\Delta \boldsymbol{y}_t^{\text{up}} = \hat{A}^T \cdot \Delta \log \boldsymbol{y}_t$.

### Industry Specific Shares:

Following Acemoglu, Akcigit, and Kerr (2016), we construct the vector of industry-specific weights by exploiting information from the input-output tables, namely: $\omega_i^{EB} = \dfrac{Sales_{i \to G}}{Sales_i}$; where "G" stands for Government.[11] By doing so, we take into account the fact that the government purchases goods and services in different quantities from each sector.[12] Lastly, the vector of weights for the EB plan, denoted by $\boldsymbol{\omega}^{EB}$, is then normalized to one.

On the contrary, we assume that aggregate TB fiscal plans impact each sector in the same fashion, therefore, we set $\omega_i^{TB} = 1/n$ for all $i$ and the $n \times 1$ vector will be: $\boldsymbol{\omega}^{TB} = 1/n \cdot \mathbf{1}_n$.

[9] see Appendix A for details on the data construction

[10] We use the Make and Use tables of year 1997, which is the closest to the occurrence of fiscal plans. Nevertheless, notice that I-O matrices are fairly stable over time.

[11] Our definition of Government encompasses both Federal and State&Local government spending. We therefore exclude here Government Enterprises, which instead are considered as part of the industrial network.

[12] We thank Roberto Perotti for this point.

# 3 Results

Equation (2) and (3) are static spatial panel autoregressive models. They allow for tracking the effect of EB and TB fiscal adjustment plans on industry output growth, controlling for downstream and upstream spillovers. Clearly they cannot be estimated by OLS since the spatial variables are endogneous. However, cutting edge spatial econometric techniques have been developed to estimate consistently this type of models. In particular, we use a modified version of the Bayesian Markov Chain Monte Carlo (MCMC) illustrated in LeSage and Pace (2009) to estimate the parameters of equation (2) and (3). In addition to the Bayesian MCMC, we also report Maximum Likelihood Estimates (MLE) for four reasons: 1. if all priors are non-informative, then the Bayesian MCMC should return the MLE, which is therefore a special case of it; 2. MLE properties of spatial panel autoregressive models with fixed effects are well known (see Yu, DeJong, and Lee (2008)); 3. it allows us to use a Vuong test to discriminate between the baseline and the inverted model. The advantage of using the Bayesian MCMC over the MLE resides in the way average aggregate effects are simulated: it is easier in the first case. Some other technical reasons make the Bayesian MCMC more appealing, however, we do not illustrate those details here, since we don't want to deviate from the macroeconomic primary focus of the paper.[13]

Table II reports descriptive statistics of the estimated parameters of interest of model (2) (baseline) and (3) (inverted); the top panel refers to the baseline model while the bottom panel refers to the inverted one. We also include year dummies for 2008 and 2009 to improve the precision of our estimates by capturing the industry-wide dip caused by the Great Recession.[14]

Firstly, looking at Table II we notice that the maximum likelihood estimates are very close to the expected value and standard deviation of the posterior distributions; a consequence of using mainly non-informative priors. Secondly, notice that during years of TB fiscal consolidations the downstream spatial correlation is much stronger than the upstream one activated during EB fiscal consolidations, in fact, $\rho^{down}$ is larger and more statistically significant (relatively smaller standard deviations) than $\rho^{up}$. Thirdly, notice that during years of TB fiscal consolidations industries' value added growth rates still exhibit stronger spatial correlation, than during years of EB fiscal consol-

[13]Nevertheless, we report in Appendix B.4 a thorough description of those technical details.

[14]For practical reasons we omit to report the estimates of the industry specific variances and fixed effects (62 times 2 estimates per model).

**Table II:** Estimation Results

| *Parameters* | MLE | | Bayesian MCMC - Posterior Distributions: | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | *Baseline Model - Equation (2)* | | | | | | | | | | | |
| | $\hat{\theta}_i^{\text{ML}}$ | MLE Std. | $\mathbb{E}(\theta_i)$ | $\sqrt{\mathbb{V}(\theta_i)}$ | $Pr(\theta_i < 0)$ | 5% | 10% | 16% | 50% | 84% | 90% | 95% |
| $\rho^{down}$ (TB) | 0.603 | 0.125 | 0.569 | 0.117 | 0.000 | 0.374 | 0.419 | 0.453 | 0.569 | 0.687 | 0.720 | 0.761 |
| $\tau_u$ | 0.411 | 1.278 | 0.555 | 1.196 | 0.322 | -1.411 | -0.971 | -0.629 | 0.551 | 1.743 | 2.095 | 2.533 |
| $\tau_a$ | -1.259 | 0.990 | -1.294 | 0.930 | 0.917 | -2.820 | -2.488 | -2.218 | -1.295 | -0.366 | -0.100 | 0.237 |
| $\tau_f$ | -0.192 | 0.432 | -0.219 | 0.404 | 0.708 | -0.887 | -0.735 | -0.621 | -0.220 | 0.182 | 0.300 | 0.447 |
| $\rho^{up}$ (EB) | 0.271 | 0.092 | 0.247 | 0.096 | 0.000 | 0.088 | 0.121 | 0.148 | 0.246 | 0.343 | 0.372 | 0.407 |
| $\gamma_u$ | -0.167 | 1.129 | -0.132 | 1.046 | 0.551 | -1.855 | -1.460 | -1.166 | -0.130 | 0.907 | 1.207 | 1.582 |
| $\gamma_a$ | 0.942 | 0.616 | 1.037 | 0.582 | 0.038 | 0.077 | 0.292 | 0.461 | 1.039 | 1.610 | 1.779 | 1.997 |
| $\gamma_f$ | -0.477 | 0.283 | -0.482 | 0.261 | 0.968 | -0.908 | -0.817 | -0.742 | -0.481 | -0.224 | -0.148 | -0.053 |
| D2008 | -2.941 | 0.671 | -2.903 | 0.633 | 1.000 | -3.946 | -3.714 | -3.532 | -2.902 | -2.274 | -2.092 | -1.861 |
| D2009 | -5.664 | 0.671 | -5.326 | 0.658 | 1.000 | -6.416 | -6.173 | -5.981 | -5.321 | -4.672 | -4.488 | -4.248 |
| | *Inverted Model - Equation (3)* | | | | | | | | | | | |
| *Parameters* | MLE | | Bayesian MCMC - Posterior Distributions: | | | | | | | | | |
| | $\hat{\theta}_i^{\text{ML}}$ | MLE Std. | $\mathbb{E}(\theta_i)$ | $\sqrt{\mathbb{V}(\theta_i)}$ | $Pr(\theta_i < 0)$ | 5% | 10% | 16% | 50% | 84% | 90% | 95% |
| $\rho^{up}$ (TB) | 0.554 | 0.103 | 0.528 | 0.097 | 0.000 | 0.368 | 0.405 | 0.432 | 0.528 | 0.625 | 0.653 | 0.687 |
| $\tau_u$ | 0.684 | 1.283 | 0.815 | 1.193 | 0.247 | -1.143 | -0.712 | -0.372 | 0.814 | 2.002 | 2.351 | 2.778 |
| $\tau_a$ | -1.298 | 0.986 | -1.290 | 0.919 | 0.920 | -2.794 | -2.463 | -2.202 | -1.293 | -0.382 | -0.112 | 0.225 |
| $\tau_f$ | -0.080 | 0.426 | -0.084 | 0.391 | 0.585 | -0.726 | -0.585 | -0.474 | -0.082 | 0.301 | 0.415 | 0.562 |
| $\rho^{down}$ (EB) | 0.096 | 0.114 | 0.125 | 0.083 | 0.000 | 0.014 | 0.026 | 0.040 | 0.112 | 0.211 | 0.241 | 0.281 |
| $\gamma_u$ | 0.073 | 1.126 | 0.050 | 1.034 | 0.480 | -1.650 | -1.272 | -0.973 | 0.051 | 1.073 | 1.370 | 1.760 |
| $\gamma_a$ | 1.286 | 0.617 | 1.296 | 0.567 | 0.011 | 0.361 | 0.572 | 0.732 | 1.295 | 1.861 | 2.023 | 2.226 |
| $\gamma_f$ | -0.502 | 0.282 | -0.499 | 0.259 | 0.973 | -0.923 | -0.831 | -0.757 | -0.499 | -0.241 | -0.169 | -0.075 |
| D2008 | -2.984 | 0.674 | -2.934 | 0.633 | 1.000 | -3.973 | -3.744 | -3.562 | -2.936 | -2.307 | -2.120 | -1.891 |
| D2009 | -5.710 | 0.674 | -5.371 | 0.661 | 1.000 | -6.469 | -6.216 | -6.025 | -5.368 | -4.717 | -4.529 | -4.290 |

Table II: *$\theta_i$ denotes a generic parameter that we estimate. The columns report the following: $\hat{\theta}_i^{ML}$ is the ML point estimate; "MLE Std." is the standard deviation of the ML estimate, calculated using the analytical Fisher Information Matrix derived in Appendix B.2: $\sqrt{\mathscr{I}(\hat{\theta}^{ML})_{ii}^{-1}}$; $\mathbb{E}(\theta_i)$ is the expected value of the posterior distribution; $\sqrt{\mathbb{V}(\theta_i)}$ is the standard deviation of the posterior distribution; $Pr(\theta < 0)$ is the probability that a parameter is negative, calculated by integrating the posterior distribution; p% is the p-th percentile of the posterior distribution. For brevity we don't report here the Industry Fixed Effects and the Industry specific variances. In the first columns, the spatial parameters also report the type of fiscal plan they are interacted with (in blue).*

idations, regardless of the inverted propagation channel. We ran a Vuong test for non-nested models adapted to our spatial model by following Wooldridge (2010).[15] Even if the results point towards a better fit of the baseline model (Equation (2)), there is not enough statistical evidence to determine its superiority over the inverted one (Equation (3)).

Concerning the fiscal coefficients, the effect of announced tax rises, $\tau_a$, and future spending cuts, $\gamma_f$, are the only ones that exhibit a strong and statistical significant recessionary effect. Interestingly, the effect of announced spending cuts,$\gamma_a$, is statistically significant and expansionary, or positive.

Nonetheless, the effect of the single coefficients of the three components of fiscal adjustment plans are not very informative: it is their convex combination into a fiscal plan that matters. Similarly, the mere size of the spatial coeffi-

[15] See Appendix B.3 for further technical details and derivation of the statistics.

cients is not enough to quantify the aggregate direct and network effect. We address these issues in the following section.

## 3.1 Aggregate Output Effect of Fiscal Consolidations

We are interested in estimating the average aggregate output effect of fiscal consolidations and then breaking it down into its direct and network effect. The spatial econometric techniques we adopt are precisely designed for such a decomposition task. The ultimate goal is to provide policymakers with reliable estimates of the output effect of TB and EB fiscal consolidations in the United States.

Firstly, fiscal consolidations are made of three components (unanticipated, anticipated and future), therefore, the standard definition of impulse response as the derivative of a dependent variable with respect to a single shock, is not applicable. What we do is to construct the impulse response by taking a convex combination of the individual derivatives of $\Delta \log \boldsymbol{y}_t$ with respect to each of the three components. Their weights will be given by the "$S$tyle" of the plan, defined analytically by:

$$\underbrace{\boldsymbol{s_{TB}}}_{3\times 1} := \left[s^u_{TB}\ s^a_{TB}\ s^f_{TB}\right]^T \qquad \underbrace{\boldsymbol{s_{EB}}}_{3\times 1} := \left[s^u_{EB}\ s^a_{EB}\ s^f_{EB}\right]^T$$

For instance, if we want to simulate a TB fiscal plan made of 30% of its unanticipated component, 0% of anticipated part and 70% of future parts, then we would set: $s^u_{TB} = .3$, $s^a_{TB} = 0$, $s^f_{TB} = .7$ and the vector of the "style" would be: $\boldsymbol{s}_{TB} = [.3\ 0\ .7]^T$.

Secondly, given: 1. the above definition of impulse response; 2. the vector representation of Equation (2); 3. the vectors of fiscal parameters $\boldsymbol{\tau}^T = [\tau_u\ \tau_a\ \tau_f]$ and $\boldsymbol{\gamma}^T = [\gamma_u\ \gamma_a\ \gamma_f]$; then, the $n \times 1$ vector of industry specific Total Effect of a TB plan ($TB_t = 1$ and $EB_t = 0$) is defined as:

$$TE_{TB} := s^u_{TB} \cdot \frac{\partial \Delta \log \boldsymbol{y}_t}{\partial f^u_t}\Big|_{TB_t=1} + s^a_{TB} \cdot \frac{\partial \Delta \log \boldsymbol{y}_t}{\partial f^a_t}\Big|_{TB_t=1} + s^f_{TB} \cdot \frac{\partial \Delta \log \boldsymbol{y}_t}{\partial f^f_t}\Big|_{TB_t=1}$$

$$= \underbrace{\left(I_n - \rho^{down} \cdot A\right)^{-1}}_{:=\boldsymbol{H}^{TB}} \cdot \boldsymbol{\omega}_{TB} \cdot \boldsymbol{\tau}^T \cdot \boldsymbol{s_{TB}} = \underbrace{\mathbf{H}^{TB} \cdot \boldsymbol{\omega}_{TB}}_{n\times 1} \cdot \underbrace{\boldsymbol{\tau}^T \cdot \boldsymbol{s_{TB}}}_{1\times 1}$$

Analogously, for an EB plan we have:

$$TE_{EB} := \underbrace{(I_n - \rho^{up} \cdot \hat{A}_0^T)^{-1}}_{:=\boldsymbol{H}^{EB}} \cdot \boldsymbol{\omega}_{EB} \cdot \boldsymbol{\gamma}^T \cdot \boldsymbol{s_{EB}} = \underbrace{\mathbf{H}^{EB} \cdot \boldsymbol{\omega}_{EB}}_{n\times 1} \cdot \underbrace{\boldsymbol{\gamma}^T \cdot \boldsymbol{s_{EB}}}_{1\times 1}.$$

Thanks to the spatial framework, the TE can be broken down into a Direct and Network Effect, as done in Acemoglu, Akcigit, and Kerr (2016) and Ozdagli and Weber (2017). The former represents the direct impact of the fiscal plan, the latter the network spillovers:

$$DE_{TB} = \boldsymbol{\omega}_{TB} \; \cdot \; \boldsymbol{\tau}^T \cdot \boldsymbol{s_{TB}} \qquad NE_{TB} = (\boldsymbol{H}^{TB} - I_n) \cdot \boldsymbol{\omega}_{TB} \; \cdot \; \boldsymbol{\tau}^T \cdot \boldsymbol{s_{TB}}$$
$$DE_{EB} = \boldsymbol{\omega}_{EB} \; \cdot \; \boldsymbol{\gamma}^T \cdot \boldsymbol{s_{EB}} \qquad NE_{EB} = (\boldsymbol{H}^{EB} - I_n) \cdot \boldsymbol{\omega}_{EB} \; \cdot \; \boldsymbol{\gamma}^T \cdot \boldsymbol{s_{EB}}$$

The TE, DE and NE are $n \times 1$ vectors of industry specific effects of fiscal adjustment plans. However, we are interested in their aggregate effect. Therefore, we take a weighted average across industries with weights given by each industry's output share.[16] By doing so we obtain the Average Total Effect, $ATE$, of a fiscal consolidation. We construct in the same way also the Average Direct Effect, $ADE$, and the Average Network Effect, $ANE$. Notice that, given the linearity of the weighted average operation, we have: $ATE = ADE + ANE$, which therefore represents a breakdown of the total effect into two components.

Table III reports descriptive statistics of the posterior distributions of ATE and its decomposition into $ADE$ and $ANE$, of 2 years long fiscal adjustment plans in the United States, which represents our main contribution to the literature of fiscal consolidations. The results are obtained via Monte-Carlo, by drawing the parameters of equation (2) and (3) from their posterior distributions.[17] The style of the simulated plans, $\boldsymbol{s}_{TB}$ and $\boldsymbol{s}_{EB}$ - which determines the composition of a fiscal plan in terms of unanticipated, anticipated, and future components - is randomly drawn at each iteration from a distribution which mimics the in-sample one and that satisfies three conditions: 1) the overall size of a plan is 1%; 2) the anticipated component is zero; 3) it has a horizon of two years.[18] By doing so, our results are robust to different styles of fiscal plans and, therefore, are not driven by a specific redistribution of the 1% fiscal shock into this or that component of the plan. This extra layer of robustness has never been used before in the fiscal consolidation literature so far.

From Table III we learn two main facts. First of all, regardless of the model analyzed and consistently with the new fact in the fiscal consolidation literature, TB fiscal consolidations imply larger output losses than EB ones.

[16] We use average output shares in years of TB fiscal consolidation for aggregating TB effects. We use average output shares in years of EB fiscal consolidation for aggregating EB effects

[17] In doing so we draw all the parameters jointly from each step of the Markov Chain to take into account the potential correlation among the parameters' distributions.

[18] See Appendix, section B.5, for further information on the empirical distribution of the style of US fiscal plans.

**Table III:** Average Total, Direct and Network Effects of Fiscal Consolidations in the United States

| | *Baseline Model - Equation (2)* | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | $\mathbb{E}(\theta)$ | *%* | $\sqrt{\mathbb{V}(\theta)}$ | $Pr(\theta < 0)$ | *5%* | *10%* | *16%* | *50%* | *84%* | *90%* | *95%* |
| $ATE_{TB}$ | -1.397 | 1 | 1.109 | 0.904 | -3.297 | -2.835 | -2.487 | -1.346 | -0.308 | -0.027 | 0.328 |
| $ADE_{TB}$ | -1.017 | 0.728 | 0.789 | 0.904 | -2.327 | -2.031 | -1.798 | -1.006 | -0.238 | -0.021 | 0.258 |
| $ANE_{TB}$ | -0.380 | 0.272 | 0.337 | 0.904 | -1.014 | -0.825 | -0.694 | -0.328 | -0.066 | -0.006 | 0.065 |
| $ATE_{EB}$ | 0.370 | 1 | 0.371 | 0.152 | -0.265 | -0.103 | 0.014 | 0.386 | 0.727 | 0.825 | 0.950 |
| $ADE_{EB}$ | 0.326 | 0.883 | 0.327 | 0.152 | -0.225 | -0.088 | 0.012 | 0.336 | 0.643 | 0.732 | 0.845 |
| $ANE_{EB}$ | 0.043 | 0.117 | 0.052 | 0.152 | -0.038 | -0.014 | 0.001 | 0.041 | 0.090 | 0.106 | 0.130 |
| | *Inverted Model - Equation (3)* | | | | | | | | | | |
| | $\mathbb{E}(\theta)$ | *%* | $\sqrt{\mathbb{V}(\theta)}$ | $Pr(\theta < 0)$ | *5%* | *10%* | *16%* | *50%* | *84%* | *90%* | *95%* |
| $ATE_{TB}$ | -1.148 | 1 | 1.034 | 0.872 | -2.909 | -2.481 | -2.162 | -1.107 | -0.131 | 0.140 | 0.480 |
| $ADE_{TB}$ | -0.848 | 0.738 | 0.756 | 0.872 | -2.106 | -1.819 | -1.593 | -0.835 | -0.101 | 0.107 | 0.375 |
| $ANE_{TB}$ | -0.300 | 0.262 | 0.290 | 0.872 | -0.828 | -0.682 | -0.572 | -0.263 | -0.029 | 0.030 | 0.102 |
| $ATE_{EB}$ | 0.522 | 1 | 0.337 | 0.064 | -0.048 | 0.096 | 0.203 | 0.536 | 0.847 | 0.936 | 1.046 |
| $ADE_{EB}$ | 0.491 | 0.940 | 0.318 | 0.064 | -0.044 | 0.089 | 0.188 | 0.501 | 0.799 | 0.886 | 0.990 |
| $ANE_{EB}$ | 0.031 | 0.060 | 0.032 | 0.064 | -0.002 | 0.002 | 0.005 | 0.024 | 0.059 | 0.073 | 0.091 |

Table III: *descriptive statistics of posterior distributions of Average Effects of a 2 years, 1% magnitude fiscal adjustment plan. 2 years means that results are calculated by cumulating the effect of the first year of the plan and then the second one. The style of the plan is simulated from a distribution which mimics the observed one; see Appendix B.4 for technical details. Columns:* $\mathbb{E}(\theta)$ *is the expected value of the posterior distribution; % is the share of ATE represented by ADE and ANE.* $\sqrt{\mathbb{V}(\theta)}$ *is the standard deviations of the posterior distribution;* $Pr(\theta < 0)$ *is the probability of negative values, calculated by integrating the posterior distribution; "p%" is the p-th percentile of the posterior distribution.*

In the baseline model the expected value of $ATE_{TB}$ is -1.397 against a positive and not significant 0.370 of $ATE_{EB}$. In the inverted model we estimate -1.148 and 0.522 for $ATE_{TB}$ and $ATE_{EB}$ respectively. This implies that a 2 years TB fiscal consolidation of 1% causes a cumulative average contraction of -1.397% over two years in the baseline model and -1.148% in the inverted one. On the other hand, the effects of EB fiscal consolidations are mildly positive and not statistically significant in the baseline model but they become positive and statistically significant in the inverted one. Given that the Vuong test does not pick any model, we consider the average -1.27% a more reliable estimate. Secondly, around 27% of $ATE_{TB}$ comes from network spillovers, confirming the relevance of the industrial network in the transmission of the TB fiscal adjustments. On the contrary, the network propagation of an EB fiscal plan is much smaller, accounting for only 11.7% of $ATE_{EB}$ in the baseline model and nearly zero in the inverted one. We calculate how much differences in the network effects of EB and TB plans account for differences in their total effects, on average: $|\mathbb{E}(ANE_{TB}) - \mathbb{E}(ANE_{EB})| \,/\, |\mathbb{E}(ATE_{TB}) - \mathbb{E}(ATE_{EB})|$. We find a value of approximately 25% in the baseline model and approximately 20% in

the inverted one.[19] We consider again the average between the baseline and inverted results as a more accurate and conservative estimate; therefore, we pick the 22.5% as the difference explained between the two.
To summarize what discovered so far: 1) we find stronger effects of TB fiscal consolidations in the United States, with an average two years contraction of around -1.27%; 2) EB fiscal consolidations in the United States are not statistically different from zero, or else, they point at mild expansionary effects after two years, in line with Alesina, Favero, and Giavazzi (2020); 3) we find evidence of network effects of TB consolidations capable of explaining 27% of the overall contraction; 4) on average, 22.5% of the differences in the ATE of TB and EB plans can be attributed to the stronger network propagation of TB fiscal consolidations.

## 3.2 Theoretical Rationalization of Empirical Findings

The results of Table III suggest that TB fiscal consolidations have a network effect which is characterized by a stronger downstream propagation: $ANE_{TB}$ is slightly larger both in levels and in share of the total effect in the baseline model, when TB plans propagate downstream. In addition to this, recall that the Vuong test suggests a very mild, but not statistically significant, superiority of the baseline model versus the inverted one. Similarly, Table III suggests that EB fiscal plans have a network effect mainly characterized by a weak upstream propagation. The estimate of $ANE_{EB}$ is weakly significant in the baseline model, when EB plans propagate upstream, and basically null in the inverted model, when EB plans propagate downstream. How do we rationalize these purely empirical findings using economic theory?

Firstly, TB fiscal consolidations are mainly characterized by pure tax increases (see Figure 2). Tax shocks can behave as both demand and supply shocks. For instance, a production or sales tax behave as pure supply shock which, according to the network literature, should propagate downstream in the production network: following a tax increase, firms reduce production and increase prices, thus hurting those customer industries which employ their good as input of production.[20] Another example is given by the negative transfers in the households budget constraint, which could have multiple effects, consistently with its promiscuous nature as both a taxing or spending device.

[19] From Table III, we have: $|-0.380 - 0.043| \,/\, |-1.397 - 0.370| \approx 25\%$ in the baseline model and $|-0.300 - 0.031| \,/\, |-1.148 - 0.522| \approx 20\%$ in the inverted model.

[20] In Appendix C we tweak the model of Acemoglu, Akcigit, and Kerr (2016) to allow for this sort of propagation mechanism using a sales tax

First, it reduces consumers' demand, which would then propagate upward in the network by reducing retailers' sales first and manufacturers' output later. Secondly, it affects labor supply, whether positively or negatively depends on the shape of the utility function and the presence of Leisure-Consumption complementarities/substitutabilities. Notice that a labor supply shift will affect wages accordingly, and then the prices of the good will incorporate such changes and finally propagate downstream in the network.

Regarding the spending side, Acemoglu, Akcigit, and Kerr (2016) model government spending shocks as pure demand shocks which propagate only upstream. Since EB fiscal consolidations are mainly made of spending cuts and have no downstream propagation in the data, our results are consistent with their theoretical predictions. Furthermore, we highlight the lack of observed downstream propagation of EB plans in the data does not necessarily mean that the EB fiscal consolidations do not trigger any such spillover. In fact, theoretically speaking the tax rise component of an EB plan could generate a mild negative downstream spillover which could be fully compensated by a positive one generated by the government spending cut. In fact, Bouakez, Rachedi, Emiliano, et al. (2018) find evidence in the United States that government spending shocks affect prices of the goods purchased by the government and then trickle down in the production network generating a downstream spillover.

Bottom line is that taxes can propagate theoretically in many directions, and this is in line with our finding that strong network effects of TB fiscal consolidations are found both in the baseline (downstream) and in the inverted model (upstream), with slightly stronger results in the former. Analogously, government spending shocks might also have complex network propagation, however, the net effect recorded in the data, within an year time span, is strongly suggesting that the propagation is weak and only upstream.

## 3.3 A Novel Macroeconometric Tool

A result of independent interest is represented by our estimation methodology. Even if we anticipated that we do not want to divert attention from the macroeconomic focus of the paper, yet, we believe certain econometric facts are worth mentioning here. This is done in the spirit of promoting the usage of these new techniques in macroeconomics.

Firstly, the adoption of spatial econometric methods allows to disentangle the direct and network effect of aggregate shocks and is a recent innovation in macroeconomics, as noted in Ozdagli and Weber (2017). Secondly, spatial models are traditionally estimated by row-normalizing the weighting matrix and removing the main diagonal from it. Another common assumption is homoskedasticity of the error term. In a recent paper Aquaro, Bailey, and Pesaran (2019) develop a new estimator which relaxes homoskedasticity and allow for different spatial coefficients (thus indirectly relaxing the row-normalization assumption); they refer to it as Heterogenous Spatial Autoregressive model (HSAR). They also point out that not assuming zero entries on the main diagonal of the weighting matrix is simply a re-parameterization of the model, which does not harm the statistical properties of the MLE, but does change the interpretation of the parameters.[21] Their econometric model, adopted by Ozdagli and Weber (2017), is very convenient for macroeconomic applications which use non-row-normalized, dense main diagonal weighting matrices and units have idiosyncratic variances. However, we highlight that even the standard dynamic spatial panel autoregressive model of Yu, DeJong, and Lee (2008), can easily be relaxed to accommodate for non-zero entries on the main diagonal, non-row-normalized weighting matrix and heteroskedastic errors, thus achieving similar results of an HSAR model; similar because HSAR allow for multiple spatial parameters.[22] The construction of a Bayesian MCMC, similar to the one in LeSage and Pace (2009), which we adopt in this paper becomes an easy and natural extension to the more general version of the spatial panel autoregressive model of Yu, DeJong, and Lee (2008). Moreover, the Bayesian MCMC method provides an easy way to recover the posterior distributions of the aggregate effects of the shocks, as illustrated earlier.

We encourage macroeconomists to adopt spatial econometric tools to study the propagation of aggregate shocks into a network of sub-units (countries, industries, regions...) but in doing so we also recommend them to follow three good practices. Firstly, always allow for heteroskedasticity, since sub-units in general have different volatilities. Secondly, never remove the main diagonal from the empirically observed weighting matrices, in our case $A$ and $\hat{A}^T$. In fact, a zero-entries main diagonal, reflects the lack in the data of spillovers within the same observed unit ("intra-unit feedback"), which is a reasonable assumption whereby units are individuals - like in standard spatial econometric applications - but it is not sensible whereby units are aggregates, such as industries. Notice, in fact, the the empirically observed $A$ and $\hat{A}^T$ weighting matrices, exhibit very dense main diagonals (see Figure 4). Thirdly, never row-normalize the weighting matrices. Row-normalization would flatten the differences in the degree of connection of each unit. For instance, in our appli-

[21] We are grateful to Hashem Pesaran for making us aware of this before the publication of their paper.

[22] We thank Lung-Fei Lee to point out to us this aspect.

cation with the industrial network, $A$ and $\hat{A}^T$ exhibit very different row-sums, indicative of different degrees of exposure to customer and supplying industries.

The bottom line is to use either the Bayesian MCMC developed here and illustrated in thorough details in Appendix B.4 or the HSAR model of Aquaro, Bailey, and Pesaran (2019) whenever the application requires different spatial coefficients. The relationship between the two models is yet to be explored in further econometric research.

# 4 Robustness

## 4.1 Spatial Model and Orders of Propagation

It could be argued that, instead of including spatial lags in the econometric model, a standard panel data model with several "cross-terms" representing the first-order, second-order, and higher-order degrees of connection should have been preferred, as in Hale, Kapan, and Minoiu (2019). However, when the network is persistent, and even higher-order propagation effects are relevant, the number of variables to be included would increase accordingly, thus increasing indefinitely the number of coefficients to estimate. On the contrary, a spatial variable is capable of capturing the entire feedback effect: an infinite number of orders of connection whose impact decays geometrically.

In order to assess whether the US industrial network with $n = 62$ sectors generates relevant high-order spillovers, we perform the partitioning of the effect similar to what suggested by LeSage and Pace (2009). For instance, for the downstream matrix we have:

$$\underbrace{(I_n - A)^{-1} \cdot \mathbf{1}_n}_{\textit{Total Effect}} = \underbrace{\mathbf{1}_n}_{\textit{Direct}} + \underbrace{A \cdot \mathbf{1}_n}_{\textit{1st order In-degree}} + \underbrace{A^2 \cdot \mathbf{1}_n}_{\textit{2nd order In-degree}} + \ldots$$

where the term in-degree refers to the fact that the row-sum of the elements of $A$ represents the weighted in-degree of the network (total share of INput purchased by a sector). For the upstream matrix, we have:

$$\underbrace{(I_n - \hat{A}^T)^{-1} \cdot \mathbf{1}_n}_{\textit{Total Effect}} = \underbrace{\mathbf{1}_n}_{\textit{Direct}} + \underbrace{\hat{A}^T \cdot \mathbf{1}_n}_{\textit{1st order Out-degree}} + \underbrace{(\hat{A}^T)^2 \cdot \mathbf{1}_n}_{\textit{2nd order Out-degree}} + \ldots$$

where the term out-degree refers to the fact that the row-sum of $\hat{A}^T$ represents the weighted out-degree of the network (total share of OUTput sold to other sectors).[23] By averaging across the 62 industries the above expressions, we can calculate how much of the average total effect (left hand side of the expressions) can be imputed to this or that order of propagation (addends of the right hand side of the expressions). The results are reported in Table IV

**Table IV:** Partitioning of the network

| *Order* | *Downstream Network* | | *Upstream Network* | |
|---|---|---|---|---|
| | *%* | *Cumulative* | *%* | *Cumulative* |
| 0 (*Direct*) | 53.36% | 53.36% | 54.53% | 54.53% |
| $1^{st}$ | 24.53% | 77.89% | 23.34% | 77.86% |
| $2^{nd}$ | 11.49% | 89.39% | 11.33% | 89.20% |
| $3^{rd}$ | 5.48% | 94.87% | 5.52% | 94.72% |
| $4^{th}$ | 2.64% | 97.51% | 2.70% | 97.42% |
| $5^{th}$ | 1.28% | 98.79% | 1.32% | 98.74% |
| $\vdots$ | $\vdots$ | $\vdots$ | $\vdots$ | $\vdots$ |

Notice that, consistently with Acemoglu, Carvalho, et al. (2012) and Carvalho (2007), the first two orders of the in-degrees and out-degrees are enough to capture most of the spillovers, roughly 89% of the overall effects. However, to capture the whole scope of network effects we should add terms up to the 5th order, which account for almost 99% of the total effect. Since we have 6 "core regressors" (TB and EB unanticipated, announced, future components), the adoption of cross terms which capture the orders of propagation, would require to include 6 times 5 orders plus one (the Direct effect) for a total of 36 core regressors. Because of this unfeasible econometric specification we opt for using a much more parsimonious spatial lag.

## 4.2 Dynamics and Delayed Network Effects

The baseline model specified by Equation (2) did not include any time lag. We adopted a fully static specification because annual industry value-added growth rates are not very persistent, in particular at the fine disaggregation level of 62 sectors. Nevertheless, few sectors still show a non-negligible degree of autocorrelation, therefore, we check whether our results are robust to the inclusion of a lagged variable. Therefore, by expressing Equation (2) (baseline)

[23] For more on in-degrees and out-degrees of the industrial network see Acemoglu, Carvalho, et al. (2012) and Carvalho and Tahbaz-Salehi (2019).

in vector notations and including a time lag, we have:

$$\underset{n\times 1}{\Delta \log \boldsymbol{y}_t} = \underset{n\times 1}{\boldsymbol{\alpha}} + \underbrace{\underset{n\times n}{\boldsymbol{\Phi}} \cdot \underset{n\times 1}{\Delta \log \boldsymbol{y}_{t-1}}}_{TimeLag}$$
$$+ \underbrace{\rho^{down} \cdot \underset{n\times n}{A} \cdot \underset{n\times 1}{\Delta \log \boldsymbol{y}_t} \cdot TB_t + \rho^{up} \cdot \underset{n\times n}{\hat{A}^T} \cdot \underset{n\times 1}{\Delta \log \boldsymbol{y}_t} \cdot EB_t}_{\text{Spatial Lag}} + \underset{n\times 6}{\boldsymbol{F}_t^T} \cdot \boldsymbol{\beta} + \nu_t \tag{4}$$

where $\boldsymbol{\Phi} = diag(\phi_1, \phi_2, ..., \phi_n)$, represents a diagonal autoregressive matrix. As done before, we also include year dummies for 2008 and 2009 to model the generalized large fall due to the Great Recession and gain precision. The results are summarized by cumulative dynamic ATE, ADE and ANE, which now take the form of cumulative impulse response functions, reported in Figure 3. The values of the median of the dynamic ATE, ADE and ANE (blue solid lines in Figure 3) are reported in Table V. Notice that after year 2, the end of the

Figure 3: Cumulative Impulse Response Functions

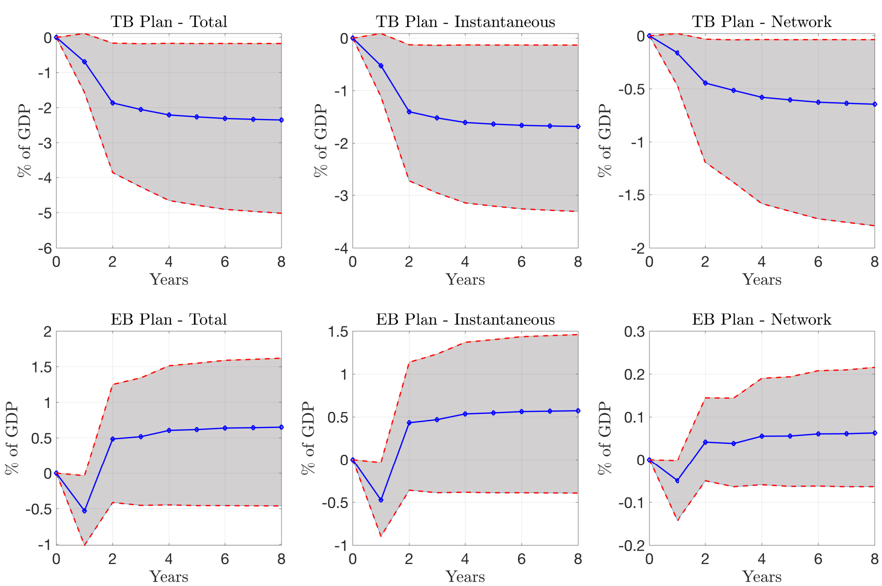



*Figure 3: blue solid lines are the median cumulative impulse response functions (median of the posterior distributions). Red dashed lines are the $5^{th}$ and $95^{th}$ percentile of their posterior distributions, which represent our confidence bands. The "shock" is constructed by simulating a two years fiscal adjustment plan of 1% of GDP, exactly as done earlier to derive our static baseline results.*

**Table V:** Median Cumulative Impulse Response Functions

| | *1 year* | % | *2 years* | % | ⋯ | *Long Run* | % |
|---|---|---|---|---|---|---|---|
| $ATE_{TB}$ | -0.695 | 100% | -1.865 | 100% | ⋯ | -2.351 | 100% |
| $ADE_{TB}$ | -0.526 | 76.7% | -1.403 | 75.2% | ⋯ | -1.683 | 71.6% |
| $ANE_{TB}$ | -0.162 | 23.3% | -0.445 | 24.8% | ⋯ | -0.644 | 28.4% |
| $ATE_{EB}$ | -0.523 | 100% | 0.486 | 100% | ⋯ | 0.628 | 100% |
| $ADE_{EB}$ | -0.472 | 90.2% | 0.433 | 89.1% | ⋯ | 0.573 | 91.2% |
| $ANE_{EB}$ | -0.049 | 9.8% | 0.041 | 10.9% | ⋯ | 0.063 | 8.8% |

fiscal consolidation, the dynamic response is minimal, corroborating our static analysis. In general, the effects are slightly larger in year 2 compared to the ones estimated in the static model and reported in Table III. Except for this, the results are comparable: 1) TB fiscal consolidations are recessionary and statistically different from zero; 2) the network effect is around one-fourth of the total effect of a TB plan; 3) EB fiscal consolidations have a minor network effect in the order of 10% of the total effect; 4) EB fiscal consolidations seem to be expansionary, but nothing can be concluded since they are not statistically different from zero.

We conclude this section by highlighting one fact: from Table V we notice that the relevance of $ANE_{TB}$ increases over time, from 23.3% to 28.4% in the long-run. This could be indicative of *delayed network effects*. Suppose a price shock takes longer than a year to travel from one sector to another, then the relevance of the network effect will increase over time since the spillover takes time to kick-in. For instance, Smets, Tielens, and Van Hove (2019) show that the autocorrelation between inflation in crude oil's price and synthetic rubber's price spikes after three months, then the autocorrelation between inflation in synthetic rubber's price and tires' price also spikes after three months but, the autocorrelation between inflation in tires' price and transport costs, spikes after 16 months. Therefore downstream propagation of price changes does seem to have delayed effects consistent with the increasing relevance over time of the network effect of TB fiscal adjustments. We stress that more research on the timing of the network effect is needed to draw any conclusion.

## 4.3 Spurious Correlation and Placebo Experiments

One result of the paper is to record significant network effects of TB fiscal consolidations, accounting for 27% of the total effect, and capable of explaining up to one fourth of the differences between the total output effect of TB and EB fiscal consolidations. What feature of the network is at basis of such strong spillovers? Are we measuring spurious correlation between sectors? or are we capturing some deep structural feature of the industrial network? We now answer these questions in the context of the baseline model, when TB fiscal consolidations propagate downstream.[24]

First of all, we plot in Figure 4 the downstream network $A$ associated with the downstream propagation of TB fiscal consolidations. Recall that the generic element of $A$, denoted by $a_{ij}$, is given by the reliance of sector $i$ (row) on industrial input $j$ (column): $SALES_{j\to i}/SALES_i$.

Figure 4: Small, medium and large elements of Downstream Network $A$


Figure 4 is a "threshold heat-map" which reports a blue cell if $a_{ij} < 0.0001$,

[24] The choice of the baseline model is motivated by its more significant effects.

an orange cell if $a_{ij} > 0.03$ and a white cell otherwise.[25] Two facts are salient from this "X-ray" of the downstream network. Firstly, the columns of $A$ tend to contain either only very small or only very large values. Secondly, the rows of $A$ do not exhibit such a pattern. In other words, some sectors, such as "Social Assistance" or "Motion Picture and Sound Recording Industries", produce an output that is either not employed at all as an intermediate by other sectors, or it is employed only in minor quantity. Unlike them, some other sectors, such as "Wholesale Trade" and "Miscellaneous Professional, Scientific and Technical Services", produce an output which is a key input of production for many sectors. The bottom line is that the US downstream network is characterized by the presence of key suppliers and the lack of key customers. This asymmetric nature of the I-O connections is a well-known feature in the production network literature (see Acemoglu, Carvalho, et al. (2012)).

An interesting robustness exercise is to see what happens to our estimates if we employ simulated network matrices that break this pattern. Basically, we estimate Equation (2) (baseline) several times by employing simulated downstream matrices ("placebo") and compare their results with the original one. We carry out two experiments:

i. *Row-Shuffling*: we randomly shuffle the order of the columns of $A$ and create 100 simulated downstream matrices. This random permutation of the columns allows us to break that natural equilibrium in which some sectors behave as key suppliers and others are marginalized. In fact, in this first simulation, some real-world key supplier might be forced to behave as a peripheral sector and vice-versa. Therefore we expect less statistically significant results.

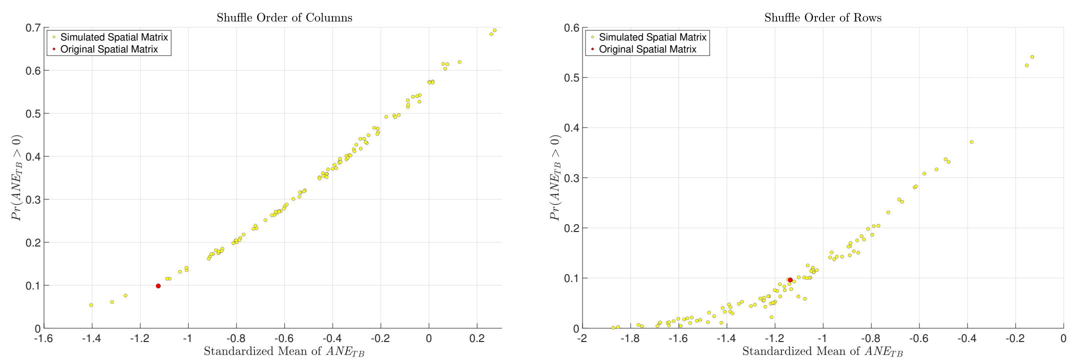
ii. *Column-Shuffling*: we randomly shuffle the order of the rows of $A$. Unlike the first experiment, reshuffling the elements within a column (shuffle the order of the rows) does not break the aforementioned characterizing pattern of the US downstream network. Sectors that originally were key suppliers will still behave in the same way. The same is true for peripheral sectors. We are reshuffling elements with similar magnitude along a column of $A$. Therefore, we expect to record both stronger and weaker results in terms of statistical significance.

Notice that in a Bayesian framework it is not fully correct to talk about statistical significance, however, with a little abuse of terminology we state that the

[25] The choice of 0.0001 is motivated by the presence of several values of $A$ which are close to zero but not exactly zero. The choice of 0.03 is motivated by the presence of only a few values above this threshold. In general, tweaking these numbers still allows observing such a visual pattern of matrix $A$.

$ANE_{TB}$ is more statistically significant if the values of $\mathbb{E}(ATE_{TB})/\sqrt{\mathbb{V}(ATE_{TB})}$ and $Pr(ATE_{TB} > 0)$ are both smaller. The first measure represents how many standard deviations we need, to obtain the average $ANE_{TB}$: the smaller it is, the more likely is to obtain sizable negative spillovers. The second measure is simply the probability of obtaining a non-negative network effect: the smaller it is, the higher the chances of getting recessionary spillovers.
Figure 5 plots on the horizontal axis the first measure and on the vertical axis the second one. The red dot represents the values obtained by employing the original matrix $A$ (see Table III). The left panel of Figure 5 reports the results

Figure 5: Placebo Experiment on $ANE_{TB}$




of the experiment of shuffling the order of the columns: the red-dot is located in the South-West region of the graph, indicative of more significant spillover effects, as expected. The right panel reports the results of the "row-shuffling" experiment: the red-dot is located almost in the middle of the cloud of simulations' results, also in line with what expected.[26]
We highlight that these three steps procedure (simulation of network matrices, re-estimation, and comparison with the original values) is analogous to Ozdagli and Weber (2017). Unlike them, our "placebo" matrices are simulated in a simpler way by simply reshuffling the orders of the columns and rows.
Our procedure has the benefit of preserving the original elements of the network matrices, thus matching one to one both the distribution of the original elements $a_{ij}$, as well as its sparsity (number of zero entries). Unlike the original network $A$, the placebo matrices do not have large entries on the main diagonal in either simulations ("dense main diagonal").

[26]Actually, slightly more dots are located more South-West than the original simulation; this is not surprising if we think that we are moving the large elements of the main diagonal (see heat-map 4) outside of it, thus mechanically inflating the indirect spillover of the sector receiving the main diagonal entry.

Concerning the first order weighted in-degrees ($A \cdot \mathbf{1}_n$) we have that the placebo matrices will exactly match it in the first simulation (shuffling the columns) while in the second one (shuffling the rows), the values are the same but they are assigned to different industries.
The second-order weighted in-degrees ($A^2 \cdot \mathbf{1}_n$) are not matched in either simulation, but the shape of their distribution is similar to the original one. Table VI summarizes the results.

**Table VI:** Placebo Experiment Results

| *Network Features:* | *Shuffling the Columns* | *Shuffling the Rows* |
|---|---|---|
| *Sparsity* | *same* | *same* |
| *Distribution of* $a_{ij}$ | *same* | *same* |
| *Dense Main Diagonal* | *no* | *no* |
| $1^{st}$ *Weighted In-degree* | *same values* | *same distribution* |
| $2^{nd}$ *Weighted In-degree* | *similar distribution* | *similar distribution* |
| *Key Suppliers* | *same* | *different* |
| *Peripheral Suppliers* | *same* | *different* |
| *Is original* $ANE_{TB}$ *stronger?* | *yes* | *no* |

Ozdagli and Weber (2017) conclude that matching the first and second order out-degree is not sufficient to justify the strong upstream propagation of monetary policy shocks. In fact, they say, matching the properties of the network industry by industry is necessary to obtain a strong network effect. We achieve the same conclusion in the context of downstream propagation of TB fiscal consolidations, measured by $ANE_{TB}$, by means of an easier experiment, namely shuffling the order of rows and columns.
Finally, we answer the initial two questions: the significant downstream network effect of TB fiscal consolidation that we find, is not capturing a spurious relationship between the sectors, otherwise its effects should not be stronger than the placebo ones when we shuffle the columns. In fact, the downstream propagation hinges on the presence of key suppliers of input of production in the industrial network, as witnessed by the lack of superior results when employing the original downstream matrix and we break this pattern (row shuffling).

# 5 Conclusion

This paper investigates the effects of fiscal consolidations and their propagation in the industrial network in the US from 1978-2014. We apply novel spatial econometric techniques to control for the spatial correlation between industries and break down the total aggregate effect of fiscal consolidations into a direct component and a network component.

Firstly, we find stronger effects of tax-based fiscal adjustments. In particular, an adjustment of one percent of GDP leads to an average contraction over two years of about -1.27% of value-added. Secondly, 27% of this effect can be attributed to spillovers from a supplying industry to a customer one. Thirdly, we find no evidence for a statistically significant recessionary impact of fiscal consolidations achieved by means of spending cuts. Rather, our evidence indicates mild expansionary effects. Fourthly, only 11% of EB effects originate from an upstream network transmission. The network effect in this case is not robust to an inverted direction of propagation: no network spillovers are detected when EB plans are let travel downstream in the network. Fifthly, we find that almost one-fourth of the different average total effects of TB and EB fiscal consolidations can be explained by stronger network spillovers of the former. Moreover, placebo experiments find that such a network effect of TB fiscal plans originates from the presence of key suppliers in the economy and does not depend on the particular shape of the distribution of first and second-order in-degrees of the network. When those key suppliers are forced to behave as peripheral suppliers the downstream propagation of TB plans vanishes or becomes significantly weaker.

Overall, our research bridges micro and macro evidence by breaking down an aggregate economy into its sectoral levels. This increases our understanding of how aggregate shocks propagate and affect the economy. In terms of policy implications, we provide further evidence that a fiscal consolidation based on spending cuts should be preferred to one based on tax hikes. The rationale is that smaller negative spillovers associated with spending cuts reduce the overall output cost. Our work has also opened the door for several extensions. For instance, our analysis reveals the likely presence of delayed network effects, which opens the door for further research on the timing of shock propagation. Also, the placebo experiments stress the importance of key suppliers of input in the industrial network. However, we do not comment on the possibility of designing optimal policies which take into account the special role of key suppliers in the propagation of shocks. We leave it to further research to address these issues.

# References


[1] Daron Acemoglu, Ufuk Akcigit, and William Kerr. “Networks and the macroeconomy: An empirical exploration.” In: *NBER Macroeconomics Annual* 30.1 (2016), pp. 273–335. DOI: 10.1086/685961.

[2] Daron Acemoglu, Vasco Carvalho, Asuman Ozdaglar, and Alireza Tahbaz-Salehi. “The network origins of aggregate fluctuations.” In: *Econometrica* 80.5 (2012), pp. 1977–2016. DOI: 10.3982/ECTA9623.

[3] Alberto Alesina, Omar Barbiero, Carlo Favero, Francesco Giavazzi, and Matteo Paradisi. “The effects of fiscal consolidations: Theory and evidence.” In: (2017). DOI: 10.3386/w23385.

[4] Alberto Alesina, Carlo Favero, and Francesco Giavazzi. *Austerity: When it Works and when it Doesn't*. Princeton University Press, 2020.

[5] Alberto Alesina, Carlo Favero, and Francesco Giavazzi. “The output effect of fiscal consolidation plans.” In: *Journal of International Economics* 96 (2015), S19–S42. DOI: 10.1016/j.jinteco.2014.11.003.

[6] Michele Aquaro, Natalia Bailey, and M Hashem Pesaran. “Estimation and inference for spatial models with heterogeneous coefficients: an application to US house prices.” In: *USC-INET Research Paper* 19-07 (2019).

[7] Alan J Auerbach, Yuriy Gorodnichenko, and Daniel Murphy. *Local Fiscal Multipliers and Fiscal Spillovers in the United States*. Tech. rep. National Bureau of Economic Research, 2019. DOI: 10.3386/w25457.

[8] David Rezza Baqaee and Emmanuel Farhi. “JEEA-FBBVA Lecture 2018: The Microeconomic Foundations of Aggregate Production Functions.” In: *Journal of the European Economic Association* 17.5 (2019), pp. 1337–1392.

[9] David Rezza Baqaee and Emmanuel Farhi. *Macroeconomics with heterogeneous agents and input-output networks*. Tech. rep. National Bureau of Economic Research, 2018.

[10] David Rezza Baqaee and Emmanuel Farhi. “The macroeconomic impact of microeconomic shocks: beyond Hulten's Theorem.” In: *Econometrica* 87.4 (2019), pp. 1155–1203.

[11] Jean-Noël Barrot and Julien Sauvagnat. “Input specificity and the propagation of idiosyncratic shocks in production networks.” In: *The Quarterly Journal of Economics* 131.3 (2016), pp. 1543–1592.

[12] Christoph E Boehm, Aaron Flaaen, and Nitya Pandalai-Nayar. "Input linkages and the transmission of shocks: Firm-level evidence from the 2011 Tōhoku earthquake." In: *Review of Economics and Statistics* 101.1 (2019), pp. 60–75.

[13] Hafedh Bouakez, Omar Rachedi, Santoro Emiliano, et al. "Sectoral heterogeneity, production networks, and the effects of government spending." In: *Cahiers de recherche. Centre interuniversitaire de recherche en economie quantitative, CIREQ* (2018).

[14] Vasco Carvalho. "Aggregate fluctuations and the network structure of intersectoral trade." In: (2007).

[15] Vasco Carvalho. "From micro to macro via production networks." In: *Journal of Economic Perspectives* 28.4 (2014), pp. 23–48. DOI: 10.1257/jep.28.4.23.

[16] Vasco Carvalho and Alireza Tahbaz-Salehi. "Production networks: A primer." In: *Annual Review of Economics* 11 (2019), pp. 635–663.

[17] Fabrice Collard, Michel Habib, and Jean-Charles Rochet. "Sovereign debt sustainability in advanced economies." In: *Journal of the European economic association* 13.3 (2015), pp. 381–420.

[18] Lydia Cox, Gernot Muller, Ernesto Pasten, Raphael Schoenle, and Michael Weber. *Big g*. Tech. rep. National Bureau of Economic Research, 2020.

[19] Xavier Gabaix. "The granular origins of aggregate fluctuations." In: *Econometrica* 79.3 (2011), pp. 733–772. DOI: 10.3982/ECTA8769.

[20] Jaime Guajardo, Daniel Leigh, and Andrea Pescatori. "Expansionary austerity? International evidence." In: *Journal of the European Economic Association* 12.4 (2014), pp. 949–968. DOI: 10.1111/jeea.12083.

[21] Galina Hale, Tumer Kapan, and Camelia Minoiu. "Shock transmission through cross-border bank lending: Credit and real effects." In: (2019). DOI: 10.17016/FEDS.2019.052.

[22] Karen J Horowitz, Mark A Planting, et al. *Concepts and Methods of the us input-Output Accounts*. Tech. rep. Bureau of Economic Analysis, 2006.

[23] LungFei Lee. "Asymptotic distributions of quasi-maximum likelihood estimators for spatial autoregressive models." In: *Econometrica* 72.6 (2004), pp. 1899–1925.

[24] James LeSage and Robert Pace. "Introduction to spatial econometrics." In: (2009). DOI: 10.1111/j.1538-4632.2010.00797.x.

[25] James LeSage and Olivier Parent. "Bayesian model averaging for spatial econometric models." In: *Geographical Analysis* 39.3 (2007), pp. 241–267. DOI: 10.1111/j.1538-4632.2007.00703.x.

[26] Christopher J Nekarda and Valerie A Ramey. "Industry evidence on the effects of government spending." In: *American Economic Journal: Macroeconomics* 3.1 (2011), pp. 36–59.

[27] Keith Ord. "Estimation methods for models of spatial interaction." In: *Journal of the American Statistical Association* 70.349 (1975), pp. 120–126. DOI: 10.1080/01621459.1975.10480272.

[28] Ali Ozdagli and Michael Weber. "Monetary policy through production networks: Evidence from the stock market." In: (2017). DOI: 10.3386/w23424.

[29] Ugo Panizza, Federico Sturzenegger, and Jeromin Zettelmeyer. "The economics and law of sovereign debt and default." In: *Journal of economic literature* 47.3 (2009), pp. 651–98.

[30] Roberto Perotti. "In search of the transmission mechanism of fiscal policy." In: *NBER Macroeconomics Annual, Chicago University Press* 22.1 (2007), pp. 169–226.

[31] Andrea Pescatori, Mr Daniel Leigh, Jaime Guajardo, and Mr Pete Devries. "A new action-based dataset of fiscal consolidation." In: 11-128 (2011). DOI: 10.5089/9781462346561.001.A001.

[32] Valerie A Ramey. "Ten years after the financial crisis: What have we learned from the renaissance in fiscal research?" In: *Journal of Economic Perspectives* 33.2 (2019), pp. 89–114.

[33] Valerie A Ramey and Matthew D Shapiro. *Costly capital reallocation and the effects of government spending*. Tech. rep. National Bureau of Economic Research, 1999.

[34] Carmen M Reinhart and Kenneth S Rogoff. *This time is different: Eight centuries of financial folly*. princeton university press, 2009.

[35] Christina D Romer and David H Romer. "The macroeconomic effects of tax changes: estimates based on a new measure of fiscal shocks." In: *American Economic Review* 100.3 (2010), pp. 763–801. DOI: 10.1257/aer.100.3.763.

[36] Frank Smets, Joris Tielens, and Jan Van Hove. "Pipeline pressures and sectoral inflation dynamics." In: (2019). DOI: 10.2139/ssrn.3346371.

[37] Quang H. Vuong. “Likelihood ratio tests for model selection and non-nested hypotheses.” In: *Econometrica* (1989). DOI: 10.2307/1912557.

[38] Jeffrey M Wooldridge. *Econometric analysis of cross section and panel data*. MIT press, 2010.

[39] Jihai Yu, Robert DeJong, and LungFei Lee. “Quasi-maximum likelihood estimators for spatial dynamic panel data with fixed effects when both n and T are large.” In: *Journal of Econometrics* 146.1 (2008), pp. 118–134.

# Appendices

## A Data on Input-Output Network

The construction of matrices $A$ and $\hat{A}^T$ starts from the analysis of the Make and Use tables illustrated in chapter 12 of Horowitz, Planting, et al. (2006).

### The Use Table

The Use table is a commodity-by-industry table which illustrates the uses of commodities by intermediate and final users. The rows of the Use Table represent the commodities (or products) and the sum of the entries in a row is the total output of that commodity. On the contrary, the columns display the industries that employ them and the final users. Horowitz, Planting, et al. (2006) provides a useful numerical example with 3 industries:

Example of Use Table - 3 Industries

| Commodity/Industry | 1 | 2 | 3 | Final demand | Total Commodity Output |
|---|---|---|---|---|---|
| 1 | 50 | 120 | 120 | 40 | 330 |
| 2 | 180 | 30 | 60 | 130 | 400 |
| 3 | 50 | 150 | 50 | 20 | 270 |
| Scrap | 1 | 3 | 1 | 0 | 5 |
| VA | 47 | 109 | 34 | / | 190 |
| Total Industry Output | 328 | 412 | 265 | 190 | / |

What is of our interest is clearly the $n \times n$ commodity-by-industry part of the Table, whose values can be denoted with the following notation:

$$(Use)_{ij} = \text{INP}_{i \to j} := \text{Commodity } i \text{ used as input by Industry } j$$

Therefore, the $n \times n$ part of the Use Table we are going to use is:

$$U = \begin{bmatrix} \text{INP}_{1\to1} & \text{INP}_{1\to2} & \text{INP}_{1\to3} \\ \\ \text{INP}_{2\to1} & \text{INP}_{2\to2} & \text{INP}_{2\to3} \\ \\ \text{INP}_{3\to1} & \text{INP}_{3\to2} & \text{INP}_{3\to3} \end{bmatrix} = \begin{bmatrix} 50 & 120 & 120 \\ 180 & 30 & 60 \\ 50 & 150 & 50 \end{bmatrix}$$

In practice, the above matrix $U$ is a "symmetric" commodity-by-industry Use Table.

Next step boils down in constructing a *commodity-by-industry direct requirement table* by dividing each industry's input, $\text{INP}_{j\to i}$, by its corresponding total industry output, $y_i$. We denote such a matrix with letter B:

$$\text{B} = \begin{bmatrix} \frac{\text{INP}_{1\to1}}{y_1} & \frac{\text{INP}_{1\to2}}{y_2} & \frac{\text{INP}_{1\to3}}{y_3} \\ \frac{\text{INP}_{2\to1}}{y_1} & \frac{\text{INP}_{2\to2}}{y_2} & \frac{\text{INP}_{2\to3}}{y_3} \\ \frac{\text{INP}_{3\to1}}{y_1} & \frac{\text{INP}_{3\to1}}{y_2} & \frac{\text{INP}_{3\to3}}{y_3} \end{bmatrix} = \begin{bmatrix} \frac{50}{328} & \frac{120}{412} & \frac{120}{265} \\ \frac{180}{328} & \frac{30}{412} & \frac{60}{265} \\ \frac{50}{328} & \frac{150}{412} & \frac{50}{265} \end{bmatrix} = \begin{bmatrix} 0.152 & 0.291 & 0.453 \\ 0.549 & 0.073 & 0.226 \\ 0.152 & 0.364 & 0.189 \end{bmatrix}.$$

Notice one important thing: matrix $B$ is different from matrix $A$, since $x_{i\to j} \neq \text{INP}_{i\to j}$: the former is an industry output flow, while the second measures a commodity flow to an industry.

### The Make Table

The Make table is an industry-by-commodity table which shows the production of commodities by industries. Row $i$ represents an industry and its summation delivers the total industry output, $y_i$. Column $j$ represents a commodity and its summation delivers the total commodity output.
Borrowing again Horowitz, Planting, et al., 2006's 3 industries example, we have:

Example of Make Table - 3 Industries

| Industry/Commodity | 1 | 2 | 3 | Scrap | Total Industry Output |
|---|---|---|---|---|---|
| 1 | 300 | 25 | 0 | 3 | 328 |
| 2 | 30 | 360 | 20 | 2 | 412 |
| 3 | 0 | 15 | 250 | 0 | 265 |
| Total Commodity Output | 330 | 400 | 270 | 5 | / |

Similarly to what done for the Use Table, we are interested in the central $n \times n$ elements of the table, which we can denote by V. The generic element of the "heart" of the Make table is:

$$(Make)_{ij} = \text{OUT}_{i\to j} := \text{Commodity } j \text{ produced by Industry } i$$

Therefore, the $n \times n$ part of the Make Table we are going to employ is:

$$V = \begin{bmatrix} \mathrm{OUT}_{1\to 1} & \mathrm{OUT}_{1\to 2} & \mathrm{OUT}_{1\to 3} \\ \\ \mathrm{OUT}_{2\to 1} & \mathrm{OUT}_{2\to 2} & \mathrm{OUT}_{2\to 3} \\ \\ \mathrm{OUT}_{3\to 1} & \mathrm{OUT}_{3\to 2} & \mathrm{OUT}_{3\to 3} \end{bmatrix} = \begin{bmatrix} 300 & 25 & 0 \\ 30 & 360 & 20 \\ 0 & 15 & 250 \end{bmatrix}$$

In practice, the above matrix $V$ is a "symmetric" industry-by-commodity Make Table.

Analogously to what done before, we now take ratios; in particular, we divide each element of V by the total production of commodity $j$. The resulting matrix is denoted by D, and its generic element is:

$$(D)_{ij} = \frac{OUT_{i\to j}}{\sum_{k=1}^{n} OUT_{k\to j}} = \frac{OUT_{i\to j}}{C_j}$$

where $C_j := \sum_{k=1}^{n} OUT_{k\to j}$ is the total production of commodity $j$. D represents the share of industry $i$ in the total production of commodity $j$; not surprisingly, Horowitz, Planting, et al. (2006) refer to this matrix as the "market share matrix". In the 3 industries/commodities example we have:

$$D = \begin{bmatrix} \frac{OUT_{1\to 1}}{C_1} & \frac{OUT_{1\to 2}}{C_2} & \frac{OUT_{1\to 3}}{C_3} \\ \\ \frac{OUT_{2\to 1}}{C_1} & \frac{OUT_{2\to 2}}{C_2} & \frac{OUT_{1\to 3}}{C_3} \\ \\ \frac{OUT_{3\to 1}}{C_1} & \frac{OUT_{3\to 2}}{C_2} & \frac{OUT_{3\to 3}}{C_3} \end{bmatrix} = \begin{bmatrix} \frac{300}{330} & \frac{25}{400} & \frac{0}{270} \\ \\ \frac{30}{330} & \frac{360}{400} & \frac{20}{270} \\ \\ \frac{0}{330} & \frac{15}{400} & \frac{250}{270} \end{bmatrix} = \begin{bmatrix} 0.909 & 0.063 & 0 \\ 0.091 & 0.900 & 0.074 \\ 0 & 0.038 & 0.926 \end{bmatrix}$$

**Adjustment for Scrap Products**

The I-O accounts include a commodity for scrap, which is a byproduct of industry production. No industry produces scrap on demand; rather, it is the result of production to meet other demands. In order to make the I-O model work correctly, we have to eliminate scrap as a secondary product. At the same time, we must also keep industry output at the same level.

This adjustment is accomplished by calculating the ratio of non-scrap output to industry output for each industry and then applying these ratios to the market shares matrix in order to account for total industry output. More precisely, the non-scrap ratio, which I denote by $\theta_i$, is defined as follows:

$$\theta_i = \frac{y_i - (\text{scrap})_i}{y_i}$$

and represents the share of total industry output $i$ made of commodity different from "scrap". In the 3 industries example we have:

| Industry | Tot.Ind.Out. | Scrap | $\Delta$ | $\theta_i$ |
|---|---|---|---|---|
| 1 | 328 | 3 | 325 | 0.991 |
| 2 | 412 | 2 | 410 | 0.995 |
| 3 | 265 | 0 | 265 | 1 |

The market shares matrix, $D$, is adjusted for scrap by dividing each row by the non-scrap ratio for that industry. In the resulting transformation matrix, called W, the implicit commodity output of each industry has been increased. In other words, we are increasing each market share to take into account that to produce each unit of each commodity, industry $i$ will produce $1/\theta_i$ units of output. In essence, we are spreading the production of commodity "scrap" over the production of all the other commodities:

$$W = \begin{bmatrix} \frac{OUT_{1\to1}}{C_1}\cdot\frac{1}{\theta_1} & \frac{OUT_{1\to2}}{C_2}\cdot\frac{1}{\theta_2} & \frac{OUT_{1\to3}}{C_3}\cdot\frac{1}{\theta_3} \\ \frac{OUT_{2\to1}}{C_1}\cdot\frac{1}{\theta_1} & \frac{OUT_{2\to2}}{C_2}\cdot\frac{1}{\theta_2} & \frac{OUT_{1\to3}}{C_3}\cdot\frac{1}{\theta_3} \\ \frac{OUT_{3\to1}}{C_1}\cdot\frac{1}{\theta_1} & \frac{OUT_{3\to2}}{C_2}\cdot\frac{1}{\theta_2} & \frac{OUT_{3\to3}}{C_3}\cdot\frac{1}{\theta_3} \end{bmatrix} = \begin{bmatrix} \frac{0.909}{0.991} & \frac{0.063}{0.991} & \frac{0}{0.991} \\ \frac{0.091}{0.995} & \frac{0.900}{0.995} & \frac{0.074}{0.995} \\ \frac{0}{1} & \frac{0.038}{1} & \frac{0.926}{1} \end{bmatrix} = \begin{bmatrix} 0.917 & 0.063 \\ 0.091 & 0.904 & 0 \\ 0 & 0.038 & 0 \end{bmatrix}$$

**The Direct Requirement Table**

To summarize:

1. We constructed matrix B, a commodity-by-industry direct requirement table, whose columns tell us how much an industry $j$ needs of commodity $i$ relative to its own total industry production.

2. We constructed matrix W, an industry-by-commodity matrix which represent the market share - adjusted for scrap - of each industry $i$ in the production of a commodity $j$.

By combining these two matrices we can obtain an industry-by-industry direct requirement matrix:

$$\underbrace{P}_{\text{industry}\times\text{industry}} := \underbrace{W}_{\text{industry}\times\text{commodity}} \cdot \underbrace{B}_{\text{commodity}\times\text{industry}}$$

In order to understand the meaning of each element of matrix P, it is important to derive it analytically:

$$P = \underbrace{\begin{bmatrix} \frac{OUT_{1\to1}}{C_1\cdot\theta_1} & \frac{OUT_{1\to2}}{C_2\cdot\theta_2} & \frac{OUT_{1\to3}}{C_3\cdot\theta_3} \\ \frac{OUT_{2\to1}}{C_1\cdot\theta_1} & \frac{OUT_{2\to2}}{C_2\cdot\theta_2} & \frac{OUT_{1\to3}}{C_3\cdot\theta_3} \\ \frac{OUT_{3\to1}}{C_1\cdot\theta_1} & \frac{OUT_{3\to2}}{C_2\cdot\theta_2} & \frac{OUT_{3\to3}}{C_3\cdot\theta_3} \end{bmatrix}}_{W} \cdot \underbrace{\begin{bmatrix} \frac{\text{INP}_{1\to1}}{y_1} & \frac{\text{INP}_{1\to2}}{y_2} & \frac{\text{INP}_{1\to3}}{y_3} \\ \frac{\text{INP}_{2\to1}}{y_1} & \frac{\text{INP}_{2\to2}}{y_2} & \frac{\text{INP}_{2\to3}}{y_3} \\ \frac{\text{INP}_{3\to1}}{y_1} & \frac{\text{INP}_{3\to2}}{y_2} & \frac{\text{INP}_{3\to3}}{y_3} \end{bmatrix}}_{B}$$

Denoting by $p_{ij}$ the generic element of P, we have:

$$p_{ij} = \frac{\frac{OUT_{i\to1}}{C_1\cdot\theta_1}\cdot\text{INP}_{1\to j} + \frac{OUT_{i\to2}}{C_2\cdot\theta_2}\cdot\text{INP}_{2\to j} + \frac{OUT_{i\to3}}{C_3\cdot\theta_3}\cdot\text{INP}_{3\to j}}{y_j} \approx \frac{\text{SALES}_{i\to j}}{SALES_j}$$

In other words, $p_{ij}$ represents how much industry $j$ depends on inputs form industry $i$ relative to its own total industry output $y_j$.[27]
Notice that the transposed of matrix P is approximately equal to matrix $A$ in the paper:

$$P \approx \begin{bmatrix} \frac{\text{SALES}_{1\to1}}{\text{SALES}_1} & \frac{\text{SALES}_{1\to2}}{\text{SALES}_2} & \frac{\text{SALES}_{1\to3}}{\text{SALES}_3} \\ \frac{\text{SALES}_{2\to1}}{\text{SALES}_1} & \frac{\text{SALES}_{2\to2}}{\text{SALES}_2} & \frac{\text{SALES}_{2\to3}}{\text{SALES}_3} \\ \frac{\text{SALES}_{3\to1}}{\text{SALES}_1} & \frac{\text{SALES}_{3\to2}}{\text{SALES}_2} & \frac{\text{SALES}_{3\to3}}{\text{SALES}_3} \end{bmatrix} \Longrightarrow \boxed{A \approx P^T}$$

[27]Notice that a big assumption is made in the construction of this matrix: if industry $i$ has adjusted market share of production of commodity $K$, $OUT_{i\to K}/(C_K\cdot\theta_K)$ equal to, say 10%, then it is assumed that if industry $j$ purchases $z := \text{INP}_{K\to j}$ dollars of commodity $K$, then 10% of $z$\$ come from industry $i$. This must be true on average but it might not be exactly true case by case.

Matrix $P$ can be either constructed from the Make and Use table or downloaded from the BEA, as an industry-by-industry direct requirement table. Its transposed value identifies the theoretical matrix $A$ in the data.
The construction of matrix $\hat{A}^T$ is trivial once we have matrix A as well as a vector of average industry output.

# B Spatial Econometric Estimation

The standard way to estimate the parameters of Equations (2) and (3) is via maximum likelihood (see LeSage and Pace (2009) for an introduction to spatial econometrics). The asymptotic and small sample properties of the MLE have been studied in Lee (2004) for cross-sectional data, and in Yu, DeJong, and Lee (2008), for dynamic panel data models with fixed effects.

## B.1 Log-likelihood

We provide here the derivation of the log-likelihood of the baseline model (2), necessary for the calculation of both the MLE and the conditional posterior distributions of the Bayesian MCMC.[28] Collecting fiscal adjustment plans, industry fixed effects and other controls into matrix $X_t$, from Equation (2):

$$
\begin{aligned}
&H_t^{-1} \cdot \underset{n\times 1}{\Delta y_t} = \underset{n\times k}{X_t} \cdot \beta + \varepsilon_t \\
&H_t = \left(I_n - \rho^{down} \cdot A \cdot TB_t - \rho^{up} \cdot \hat{A}^T \cdot EB_t\right)^{-1} \\
&\varepsilon_t \sim \mathscr{N}(0, \Omega), \forall t \in \{1, ..., T\} \\
&\Omega = diag(\sigma_1^2, ..., \sigma_n^2) \\
&\varepsilon_t \perp \varepsilon_{t+i}, \quad \forall t \in \{1, ..., T\}, \forall i \in \mathscr{Z}
\end{aligned}
$$

where $k$ is the number of regressors.[29] We now make a convenient change in the notation: 1. we now use $'$ as a symbol for transposition instead of $^T$; 2. we now set $\rho_1 = \rho^{down}$, $\rho_2 = \rho^{up}$, $A = W_1$ and $\hat{A}' = W_2$. We have:

$$
Z_t := H_t^{-1} \cdot \Delta y_t \sim \mathscr{N}(X_t\beta, \Omega) \implies \Delta y_t \sim \mathscr{N}(H_t X_t \beta, H_t \Omega H_t')
$$

The density function of the random vector $\Delta y_t$ is:

$$
f(\underset{n\times 1}{\Delta y_t} | X_t, \rho, \beta, \Omega) = \frac{1}{\sqrt{(2\pi)^n \cdot |H_t \Omega H_t'|}} \exp\left\{-\frac{1}{2} \cdot (\Delta y_t - H_t X_t \beta)' \cdot (H_t \Omega H_t')^{-1} \cdot (\Delta y_t - H_t X_t \beta)\right\},
$$

[28] Results for the inverted model, Equation (3)) are symmetric to the baseline case.

[29] $k$ in our baseline is $n$ fixed effects plust 6 fiscal adjustment components (unexpected, announced and future for both TB and EB plans) plus 2 year dummies for 2008 and 2009.

with $\rho = \left[\rho^{down}, \rho^{up}\right]$.
Given that $(H_t \Omega H_t')^{-1} = (H_t')^{-1} \cdot \Omega^{-1} \cdot H_t^{-1}$ and $|H_t \Omega H_t'| = |H_t|^2 \cdot |\Omega|$, we have:

$$\begin{aligned}
f(\Delta y_t|\cdot) &= (2\pi)^{-n/2} \cdot |H_t|^{-1} \cdot |\Omega|^{-1/2} \cdot \exp\Bigg\{ -\frac{1}{2}(Z_t - X_t\beta)' \cdot H_t' \cdot (H_t')^{-1} \cdot \Omega^{-1} \cdot H_t^{-1} \cdot H_t \cdot (Z_t - X_t \\
&= (2\pi)^{-n/2} \cdot |(I_n - \rho_1 W1TB_t - \rho_2 W_2 EB_t)^{-1}|^{-1} \cdot |\Omega|^{-1/2} \exp\Bigg\{ -\frac{1}{2}\varepsilon_t'\Omega^{-1}\varepsilon_t \Bigg\} \\
&= (2\pi)^{-n/2} \cdot |I_n - \rho_1 \cdot W_1 \cdot TB_t - \rho_2 \cdot W_2 \cdot EB_t| \cdot |\Omega|^{-1/2} \exp\Bigg\{ -\frac{1}{2}\varepsilon_t'\Omega^{-1}\varepsilon_t \Bigg\},
\end{aligned}$$

At this point we need to find the likelihood of the random vector $\Delta y = \begin{bmatrix} \Delta y_1' & \dots & \Delta y_T' \end{bmatrix}$. Since the model is static and we have assumed $cov(\varepsilon_t, \varepsilon_{t-k}) = \underset{n\times n}{\rho^0}$, then $\Delta y_t$ is *iid* over time. By consequence, the following holds:

$$f(\underset{nT\times 1}{\Delta y}|X_1, \dots, X_T, \rho, \beta, \Omega) = \prod_{t=1}^{T} f(\underset{n\times 1}{\Delta y_t}|X_t, \rho, \beta, \Omega) = \big((2\pi)^n|\Omega|\big)^{-T/2} \cdot$$

$$\cdot \prod_{t=1}^{T} |I_n - \rho_1 \cdot W_1 \cdot TB_t - \rho_2 \cdot W_2 \cdot EB_t| \exp\big\{ -\frac{1}{2} \cdot \sum_{t=1}^{T} \varepsilon_t'\Omega^{-1}\varepsilon_t \big\}.$$

Now we divide the time series of length $T$ in three different sub-periods. In doing so, consider the following new parameters:

- $t_1$: set of years when a tax based fiscal adjustment occurs.
  Formally $t_1 := \{1, ..., t, ..., T_1|\ t \text{ such that } TB_t = 1\}$. We set: $H_t|\ t \in t_1 = (I_n - \rho_1 \cdot W_1)^{-1} = H_\tau$.

- $t_2$: set of years when an expenditure tax based fiscal adjustment occurs.
  Formally: $t_2 := \{1, ..., t, ..., T_2|\ t \text{ such that } EB_t = 1\}$. We set $H_t|\ t \in t_2 = (I_n - \rho_2 \cdot W_2)^{-1} = H_\gamma$.

- $t_3$: set of years when neither a tax based fiscal adjustment nor an expenditure based fiscal adjustment occurs.
  Formally $t_3 := \{1, ..., t, ..., T_3|\ t \text{ such that } TB_t = 0 \wedge EB_t = 0\}$. We set $H_t|\ t \in t_3 = (I_n)^{-1} = I_n$.

Therefore, we have that $t_1$, $t_2$ and $t_3$ account for a partition of the whole time series and $T = T_1 + T_2 + T_3$. By consequence we have:

$$
\begin{aligned}
\prod_{t=1}^{T} |I_n - \rho_1 W_1 TB_t - \rho_2 W_2 EB_t| &= \prod_{t=1}^{T} |H_t^{-1}| \\
&= \prod_{t=1}^{T} \frac{1}{|H_t|} \\
&= \prod_{t\in t_1}^{T_1} \frac{1}{|H_t|} \cdot \prod_{t\in t_2}^{T_2} \frac{1}{|H_t|} \cdot \prod_{t\in t_3}^{T_3} \frac{1}{|H_t|} \\
&= |H_\tau|^{-T_1} \cdot |H_\gamma|^{-T_2} \cdot |I_n|^{-T_3} \\
&= |I_n - \rho_1 \cdot W_1|^{T_1} \cdot |I_n - \rho_2 W_2|^{T_2}
\end{aligned}
$$

At this point, we rewrite the probability density function of our dependent variable as:

$$
\begin{aligned}
f(\Delta y_t | X_1, \ldots, X_T, \rho, \beta, \Omega) &= (2\pi)^{-nT/2} \cdot |\Omega|^{-T/2} \cdot \\
&\cdot |I_n - \rho_1 \cdot W_1|^{T_1} \cdot |I_n - \rho_2 W_2|^{T_2} \cdot \exp\left\{ -\frac{1}{2} \cdot \sum_{t=1}^{T} \varepsilon_t' \cdot \Omega^{-1} \cdot \varepsilon_t \right\}.
\end{aligned}
$$

Finally, the log-likelihood of our dataset is:

$$
\begin{aligned}
\log \mathscr{L}(\rho, \beta, \Omega | \Delta y_1, \ldots, \Delta y_T, X_1, \ldots, X_T) &= -\frac{nT}{2} \ln(2\pi) - \frac{T}{2} \cdot \ln(|\Omega|) + \\
&+ T_1 \cdot \ln(|I_n - \rho_1 \cdot W_1|) + T_2 \cdot \ln(|I_n - \rho_2 W_2|) - \frac{1}{2} \cdot \sum_{t=1}^{T} \varepsilon_t' \cdot \Omega^{-1} \cdot \varepsilon_t.
\end{aligned}
$$

with:

$$
\varepsilon_t = Z_t - X_t \cdot \beta = H_t^{-1} \cdot \Delta y_t - X_t \beta = (I_n - \rho_1 W_1 TB_t - \rho_2 W_2 EB_t) \cdot \Delta y_t - X_t \cdot \beta.
$$

Furthermore, we impose the condition $\lambda_{\min}^{-1} < \hat{\rho}_1 < \lambda_{\max}^{-1}$ and $\mu_{\min}^{-1} < \hat{\rho}_2 < \mu_{\max}^{-1}$, where $\lambda$ and $\mu$ are the eigenvalues of the spatial matrices $W_1$ and $W_2$ respectively. This condition guarantees that the estimated model will have positive definite covariance matrix (see Ord (1975)).
Notice that in the inverted model of Equation (3), it is enough to switch the definition of $W_1$ and $W_2$ by setting: $A = W_2$ and $\hat{A}' = W_1$.

## B.2 The Analytical Fisher Information Matrix

In order to derive the Fisher Information Matrix we firstly need to obtain the total gradient of the log-likelihood function. Let's start with the spatial coefficient $\rho_1$:

$$\frac{\partial \log \mathscr{L}(\theta|\Delta y, X)}{\partial \rho_1} = T_1 \frac{1}{|I_n - \rho_1 W_1|} \frac{\partial |I_n - \rho_1 W_1|}{\partial \rho_1} - \frac{1}{2} \sum_{t=1}^{T} \frac{\partial (Z_t' \Omega^{-1} Z_t)}{\partial \rho_1} - 2 \frac{\partial (Z_t' \Omega^{-1} X_t \beta)}{\partial \rho_1}.$$

By some matrix algebra, it is possible to show that:

$$\begin{aligned} \frac{\partial (Z_t' \Omega^{-1} Z_t)}{\partial \rho_1} &= -TB_t \cdot \Delta y_t' \cdot \Omega^{-1} \cdot W_1 \cdot \Delta y_t - TB_t \cdot \Delta y_t' \cdot W_1' \Omega^{-1} \cdot \Delta y_t \\ &+ 2\rho_1 \cdot TB_t^2 \cdot \Delta y_t' \cdot W_1 \cdot \Omega^{-1} \cdot W_1 \cdot \Delta y_t' + 2\rho_2 \cdot TB_t \cdot EB_t \cdot \Delta y_t' \cdot W_1 \cdot \Omega^{-1} \cdot W_2 \cdot \Delta y_t' \end{aligned}$$

Since our fiscal adjustment plans are mutually exclusive, we have that $TB_t \cdot EB_t = 0$ for all $t$. Moreover, by rearranging the above expression, we get:

$$\frac{\partial (Z_t' \Omega^{-1} Z_t)}{\partial \rho_1} = -2 \cdot TB_t \cdot \Delta y_t' \cdot (I_n - \rho_1 \cdot W_1') \cdot \Omega^{-1} \cdot W_1 \cdot \Delta y_t$$

After other matrix algebra, we get:

$$-2 \cdot \frac{\partial (Z_t \cdot \Omega^{-1} X_t \beta)}{\partial \rho_1} = 2 \cdot TB_t \cdot \Delta y_t' \cdot W_1' \cdot \Omega^{-1} \cdot X_t \cdot \beta$$

Wrapping up all together, and employing the notation introduced earlier: $(I_n - \rho_1 W_1)^{-1} = H_\tau$, we have:

$$\begin{aligned} \frac{\partial \log \mathscr{L}(\theta|\Delta y, X)}{\partial \rho_1} &= T_1 \frac{1}{|I_n - \rho_1 W_1|} \frac{\partial |I_n - \rho_1 W_1|}{\partial \rho_1} + \\ &+ \sum_{t \in t_1}^{T_1} \Big[ \Delta y_t' \cdot (I_n - \rho_1 \cdot W_1') \cdot \Omega^{-1} \cdot W_1 \cdot \Delta y_t - \Delta y_t' \cdot W_1' \cdot \Omega^{-1} \cdot X_t \cdot \beta \Big] = \\ &= T_1 \frac{1}{|I_n - \rho_1 W_1|} \cdot |I_n - \rho_1 W_1| \cdot Tr\Big( (I_n - \rho_1 W_1)^{-1} \cdot (-W_1) \Big) + \\ &+ \sum_{t \in t_1}^{T_1} \Big[ \big( (I_n - \rho_1 \cdot W_1) \cdot \Delta y_t \big)' \cdot \Omega^{-1} \cdot W_1 \cdot \Delta y_t - \beta' \cdot X_t' \cdot \Omega^{-1} \cdot W_1 \cdot \Delta y_t \Big] \\ &= -T_1 \cdot Tr\Big( H_\tau \cdot W_1 \Big) + \sum_{t \in t_1}^{T_1} \Big[ \big( Z_t - X_t \beta \big)' \cdot \Omega^{-1} \cdot W_1 \cdot \Delta y_t \Big] \\ &= \sum_{t \in t_1}^{T_1} \big( \varepsilon_t' \cdot \Omega^{-1} \cdot W_1 \cdot \Delta y_t \big) - T_1 \cdot Tr\big( H_\tau \cdot W_1 \big). \end{aligned}$$

By simmetry we have that:

$$\frac{\partial \log \mathscr{L}(\theta|\Delta y, X)}{\partial \rho_2} = \sum_{t \in t_2}^{T_2} \big(\varepsilon_t' \cdot \Omega^{-1} \cdot W_2 \cdot \Delta y_t\big) - T_2 \cdot Tr\Big(H_\gamma \cdot W_2\Big),$$

with $H_\gamma = (I_n - \rho_2 W_2)^{-1}$, from the previous notation.
As far as concern the derivative with respect to $\beta$, we have already seen when concentrating the log-likelihood that:

$$\begin{aligned}
\frac{\partial \log \mathscr{L}(\theta|\Delta y, X)}{\partial \beta} &= X' \cdot \Sigma^{-1} \cdot Z - X' \cdot \Sigma^{-1} \cdot X \cdot \beta \\
&= X' \cdot \Sigma^{-1} \cdot (Z - X \cdot \beta) = \\
&= X' \cdot \Sigma^{-1} \cdot \varepsilon = \\
&= \sum_{t=1}^{T} X_t' \cdot \Omega^{-1} \cdot \varepsilon_t.
\end{aligned}$$

Concerning the derivatives with respect to $\sigma_i^2$, we need firstly to acknowledge that:

$$\sum_{t=1}^{T} \varepsilon_t' \cdot \Omega^{-1} \cdot \varepsilon_t = \sum_{t=1}^{T} \sum_{i=1}^{n} \frac{\varepsilon_{i,t}^2}{\sigma_i^2} = \sum_{i=1}^{n} \frac{1}{\sigma_i^2} \sum_{t=1}^{T} \varepsilon_{i,t}^2,$$

and that:

$$\ln(|\Omega|) = \ln(\prod_{i=1}^{n} \sigma_i^2) = \sum_{i=1}^{n} \ln(\sigma_i^2).$$

Therefore, we have that:

$$\begin{aligned}
\frac{\partial \log \mathscr{L}(\theta|\Delta y, X)}{\partial \sigma_i^2} &= -\frac{T}{2} \frac{\partial \ln(|\Omega|)}{\partial \sigma_i^2} - \frac{1}{2} \cdot \frac{\partial}{\partial \sigma_i^2} \sum_{t=1}^{T} \varepsilon_t' \cdot \Omega^{-1} \cdot \varepsilon_t \\
&= -\frac{T}{2 \cdot \sigma_i^2} + \frac{1}{2 \cdot \sigma_i^4} \cdot \sum_{t=1}^{T} \varepsilon_{i,t}^2.
\end{aligned}$$

We now have all the elements to write down the gradient of the log-likelihood:

$$\nabla \log \mathscr{L}(\theta|\Delta y, X) = \underbrace{\begin{bmatrix} \dfrac{\partial \log \mathscr{L}(\theta|\Delta y, X)}{\partial \rho_1} \\ \\ \dfrac{\partial \log \mathscr{L}(\theta|\Delta y, X)}{\partial \rho_2} \\ \\ \dfrac{\partial \log \mathscr{L}(\theta|\Delta y, X)}{\partial \beta} \\ \\ \dfrac{\partial \log \mathscr{L}(\theta|\Delta y, X)}{\partial \sigma_1^2} \\ \\ \vdots \\ \\ \dfrac{\partial \log \mathscr{L}(\theta|\Delta y, X)}{\partial \sigma_n^2} \end{bmatrix}}_{38\times 1} = \begin{bmatrix} \sum_{t\in t_1}^{T_1} \big(\varepsilon_t' \cdot \Omega^{-1} \cdot W_1 \cdot \Delta y_t\big) - T_1 \cdot Tr\big(H_\tau \cdot W_1\big) \\ \\ \sum_{t\in t_2}^{T_2} \big(\varepsilon_t' \cdot \Omega^{-1} \cdot W_2 \cdot \Delta y_t\big) - T_2 \cdot Tr\big(H_\gamma \cdot W_2\big) \\ \\ \sum_{t=1}^{T} X_t' \cdot \Omega^{-1} \cdot \varepsilon_t \\ \\ -\dfrac{T}{2\cdot\sigma_1^2} + \dfrac{1}{2\cdot\sigma_1^4} \cdot \sum_{t=1}^{T} \varepsilon_{1,t}^2 \\ \\ \vdots \\ \\ -\dfrac{T}{2\cdot\sigma_n^2} + \dfrac{1}{2\cdot\sigma_n^4} \cdot \sum_{t=1}^{T} \varepsilon_{n,t}^2 \end{bmatrix}$$

Another round of derivation is now needed. Let's start with the first row of the matrix: all the derivatives of $\dfrac{\partial \log \mathscr{L}(\theta|\Delta y, X)}{\partial \rho_1}$ with respect to all the parameters. To simplify notation we will refer with $\mathscr{H}_{ij}$ to the element of row $i$ and column $j$ of the Hessian matrix.

$$\begin{aligned}
\mathscr{H}_{1,1} = \frac{\partial^2 \log \mathscr{L}(\theta|\Delta y, X)}{\partial \rho_1^2} &= \sum_{t\in t_1}^{T_1} \big(\frac{\partial \varepsilon_t'}{\partial \rho_1} \cdot \Omega^{-1} \cdot W_1 \cdot \Delta y_t\big) - T_1 \cdot \frac{\partial Tr\big(H_\tau \cdot W_1\big)}{\partial \rho_1} \\
&= \sum_{t\in t_1}^{T_1} \big((-\Delta y_t' \cdot W_1') \cdot \Omega^{-1} \cdot W_1 \cdot \Delta y_t\big) - T_1 \cdot Tr\big(\frac{\partial H_\tau}{\partial \rho_1} \cdot W_1\big) = \\
&= -\sum_{t\in t_1}^{T_1} \big(\Delta y_t' \cdot W_1' \cdot \Omega^{-1} \cdot W_1 \cdot \Delta y_t\big) - T_1 \cdot Tr\big((-H_\tau \cdot (-W_1) \cdot H_\tau) \cdot W_1\big) = \\
&= -T_1 \cdot Tr\big(W_1 \cdot H_\tau \cdot W_1 \cdot H_\tau\big) - \sum_{t\in t_1}^{T_1} \big(\Delta y_t' \cdot W_1' \cdot \Omega^{-1} \cdot W_1 \cdot \Delta y_t\big)
\end{aligned}$$

Symmetrically we have:

$$\mathscr{H}_{2,2} = \frac{\partial^2 \log \mathscr{L}(\theta|\Delta y, X)}{\partial \rho_2^2} =$$
$$= -T_2 \cdot Tr\big(W_2 \cdot H_\gamma \cdot W_2 \cdot H_\gamma\big) - \sum_{t\in t_2}^{T_2} \big(\Delta y_t' \cdot W_2' \cdot \Omega^{-1} \cdot W_2 \cdot \Delta y_t\big)$$

Going back to the first row, we now calculate the cross derivative with respect to $\rho 2$. Before doing so, recall that, being the log-likelihood a continuously diffirentiable function, the Schwarz's theorem applies and the Hessian matrix is symmetric.

$$\mathscr{H}_{1,2} = \mathscr{H}_{2,1} = \frac{\partial^2 \log \mathscr{L}(\theta|\Delta y, X)}{\partial \rho_1 \partial \rho_2} = 0.$$

Going on with the calculation we have:

$$\begin{aligned}
\mathscr{H}_{1,3:1,23} = \frac{\partial^2 \log \mathscr{L}(\theta|\Delta y, X)}{\partial \rho_1 \partial \beta} &= \sum_{t\in t_1}^{T_1} \big(\frac{\partial \varepsilon_t'}{\partial \beta} \cdot \Omega^{-1} \cdot W_1 \cdot \Delta y_t\big) \\
&= -\sum_{t\in t_1}^{T_1} X_t' \cdot \Omega^{-1} \cdot W_1 \cdot \Delta y_t \\
&= -X_\tau' \cdot \underbrace{(I_{T_1} \otimes \Omega^{-1})}_{\Sigma_\tau^{-1}} \cdot (I_{T_1} \otimes W_1) \cdot \Delta y_\tau
\end{aligned}$$

where $\mathscr{H}_{1,3:1,23}$ means all the elements of the first row, from column 3 up to column 23. $X_\tau$ and $\Delta y_\tau$ represent $X$ and $\Delta y$ but for the only years when a tax based fiscal adjustment occur:

$$X_\tau = \underset{T_1 n\times k}{\begin{bmatrix} X_1 \\ \vdots \\ X_t \\ \vdots \\ X_{T_1} \end{bmatrix}} \quad \text{and} \quad \Delta y_\tau = \underset{T_1 n\times k}{\begin{bmatrix} \Delta y_1 \\ \vdots \\ \Delta y_t \\ \vdots \\ \Delta y_{T_1} \end{bmatrix}} \quad \text{with } t \in t_1,$$

Symmetrically:

$$\begin{aligned}\mathscr{H}_{2,3:2,23} = \frac{\partial^2 \log \mathscr{L}(\theta|\Delta y, X)}{\partial \rho_2 \partial \beta} &= \sum_{t \in t_2}^{T_2} \big( \frac{\partial \varepsilon_t'}{\partial \beta} \cdot \Omega^{-1} \cdot W_2 \cdot \Delta y_t \big) \\ &= -\sum_{t \in t_2}^{T_2} X_t' \cdot \Omega^{-1} \cdot W_2 \cdot \Delta y_t \\ &= -X_\gamma' \cdot (\underbrace{I_{T_2} \otimes \Omega^{-1}}_{\Sigma_\gamma^{-1}}) \cdot (I_{T_2} \otimes W_2) \cdot \Delta y_\gamma,\end{aligned}$$

with:

$$X_\gamma = \underset{T_2 n \times k}{\begin{bmatrix} X_1 \\ \vdots \\ X_t \\ \vdots \\ X_{T_2} \end{bmatrix}} \quad \text{and} \quad \Delta y_\gamma = \underset{T_2 n \times k}{\begin{bmatrix} \Delta y_1 \\ \vdots \\ \Delta y_t \\ \vdots \\ \Delta y_{T_2} \end{bmatrix}} \quad \text{with } t \in t_2,$$

$$\begin{aligned}\mathscr{H}_{3,3:23,23} = \frac{\partial^2 \log \mathscr{L}(\theta|\Delta y, X)}{\partial \beta^2} &= \frac{\partial}{\partial \beta^2} \left( \sum_{t=1}^{T} X_t' \cdot \Omega^{-1} \cdot \varepsilon_t \right) \\ &= \sum_{t=1}^{T} X_t' \cdot \Omega^{-1} \cdot \frac{\partial (Z_t - X_t \cdot \beta)}{\partial \beta^2} \\ &= \sum_{t=1}^{T} X_t' \cdot \Omega^{-1} \cdot X_t \\ &= -X' \cdot \Sigma^{-1} \cdot X.\end{aligned}$$

$$\mathscr{H}_{3,24:23,38} = \frac{\partial^2 \log \mathscr{L}(\theta|\Delta y, X)}{\partial \beta \partial \sigma^2} = \sum_{t=1}^{T} X_t' \cdot \frac{\partial \Omega^{-1}}{\partial \sigma^2} \cdot \varepsilon_t$$

The generic element of the above matrix is a $k \times 1$ vector:

$$-\sigma_1^{-4} \cdot \sum_{t=1}^{T} X_{1,t}' \cdot \varepsilon_{i,t}.$$

Going on with the calculation:

$$\mathscr{H}_{i,i|i\in[24,38]} = \frac{\partial^2 \log \mathscr{L}(\theta|\Delta y, X)}{\partial(\sigma_i^2)^2} = \frac{T}{2}\cdot\frac{1}{\sigma_i^4}\cdot\Big(1-\frac{2}{T\cdot\sigma_i^2}\cdot\sum_{t=1}^{T}\varepsilon_{i,t}^2\Big).$$

$$\mathscr{H}_{23+i,23+j|i,j\in[1,n]} = \frac{\partial^2 \log \mathscr{L}(\theta|\Delta y, X)}{\partial\sigma_i^2\partial\sigma_j^2} = 0 \quad \forall i \neq j.$$

$$\begin{aligned}
\mathscr{H}_{1,24:1,38} = \frac{\partial^2 \log \mathscr{L}(\theta|\Delta y, X)}{\partial\rho_1\partial\sigma_i^2} &= \frac{\partial}{\partial\sigma_i^2}\Big(\sum_{t\in t_1}^{T_1}\varepsilon_t'\cdot\Omega^{-1}\cdot W_1\cdot\Delta y_t\Big)\\
&= \frac{\partial}{\partial\sigma_i^2}\Big(\sum_{t\in t_1}^{T_1} Tr\big(\varepsilon_t'\cdot\Omega^{-1}\cdot W_1\cdot\Delta y_t\big)\Big)\\
&= \frac{\partial}{\partial\sigma_i^2}\Big(Tr\big((\sum_{t\in t_1}^{T_1}\Delta y_t\cdot\varepsilon_t')\cdot\Omega^{-1}\cdot W_1\big)\Big)\\
&= Tr\Big((\sum_{t\in t_1}^{T_1}\Delta y_t\cdot\varepsilon_t')\cdot\frac{\partial\Omega^{-1}}{\partial\sigma_i^2}\cdot W_1\Big)
\end{aligned}$$

Note that

$$\frac{\partial\Omega^{-1}}{\partial\sigma_i^2} = \begin{bmatrix} 0 & \cdots & 0 & \cdots & 0\\ \vdots & \ddots & \vdots & & \vdots\\ 0 & \cdots & -\sigma_i^{-4} & \cdots & 0\\ \vdots & & \vdots & \ddots & \vdots\\ 0 & \cdots & 0 & \cdots & 0\end{bmatrix} = diag(0,\cdots,0,-\sigma_i^{-4},0,\cdots,0)$$

Symmetrically:

$$\mathscr{H}_{2,24:2,38} = \frac{\partial^2 \log \mathscr{L}(\theta|\Delta y, X)}{\partial\rho_2\partial\sigma_i^2} = Tr\Big((\sum_{t\in t_2}^{T_2}\Delta y_t\cdot\varepsilon_t')\cdot\frac{\partial\Omega^{-1}}{\partial\sigma_i^2}\cdot W_2\Big)$$

At this point we have all the elements to construct the Hessian matrix of the log-likelihood.
To sum up, first row:

- $\mathscr{H}_{1,1} = -T_1\cdot Tr\big(W_1\cdot H_\tau\cdot W_1\cdot H_\tau\big) - \sum_{t\in t_1}^{T_1}\big(\Delta y_t'\cdot W_1'\cdot\Omega^{-1}\cdot W_1\cdot\Delta y_t\big)$

- $\mathscr{H}_{1,2} = 0$

- $\mathscr{H}_{1,3:1,23} = -\sum_{t\in t_1}^{T_1} X_t' \cdot \Omega^{-1} \cdot W_1 \cdot \Delta y_t$

- $\mathscr{H}_{1,24:1,38} = Tr\Big(\big(\sum_{t\in t_1}^{T_1} \Delta y_t \cdot \varepsilon_t'\big) \cdot \dfrac{\partial\Omega^{-1}}{\partial\sigma_i^2} \cdot W_1\Big).$

Second row:

- $\mathscr{H}_{2,1} = 0$

- $\mathscr{H}_{2,2} = -T_2 \cdot Tr\big(W_2 \cdot H_\gamma \cdot W_2 \cdot H_\gamma\big) - \sum_{t\in t_2}^{T_2} \big(\Delta y_t' \cdot W_2' \cdot \Omega^{-1} \cdot W_2 \cdot \Delta y_t\big)$

- $\mathscr{H}_{2,3:2,23} = -\sum_{t\in t_2}^{T_2} X_t' \cdot \Omega^{-1} \cdot W_2 \cdot \Delta y_t$

- $\mathscr{H}_{2,24:2,38} = Tr\Big(\big(\sum_{t\in t_2}^{T_2} \Delta y_t \cdot \varepsilon_t'\big) \cdot \dfrac{\partial\Omega^{-1}}{\partial\sigma_i^2} \cdot W_2\Big).$

From row 3 to row 23:

- $\mathscr{H}_{3,1:23,1} = \mathscr{H}_{1,3:1,23}'$

- $\mathscr{H}_{3,2:23,2} = \mathscr{H}_{2,3:2,23}'$

- $\mathscr{H}_{3,3:23,23} = \sum_{t=1}^{T} X_t' \cdot \Omega^{-1} \cdot X_t$

- $\mathscr{H}_{3,24:23,38} = \sum_{t=1}^{T} X_t' \cdot \dfrac{\partial\Omega^{-1}}{\partial\sigma^2} \cdot \varepsilon_t$

From row 24 to the last row (number 38):

- $\mathscr{H}_{24,1:38,1} = \mathscr{H}_{1,24:1,38}'$

- $\mathscr{H}_{24,2:38,2} = \mathscr{H}_{2,24:2,38}'$

- $\mathscr{H}_{24,3:38,23} = \mathscr{H}_{3,24:23,38}'$

- $\mathscr{H}_{23+i,23+j|i,j\in[1,n]} = \begin{cases} \dfrac{T}{2}\cdot\dfrac{1}{\sigma_i^4}\cdot\left(1-\dfrac{2}{T\cdot\sigma_i^2}\cdot\displaystyle\sum_{t=1}^{T}\varepsilon_{i,t}^2\right) & \forall i=j\in[1,n] \\ \\ 0 \quad \forall i\neq j \end{cases}$

The last step we have to make to finally obtain the Fisher Information Matrix is taking expectations of every element.

$$
\begin{aligned}
E[\mathscr{H}_{1,1}] &= -T_1\cdot Tr\big(W_1\cdot H_\tau\cdot W_1\cdot H_\tau\big) - \sum_{t\in t_1}^{T_1} E\big[\Delta y_t'\cdot W_1'\cdot\Omega^{-1}\cdot W_1\cdot\Delta y_t\big] = \\
&= -T_1\cdot Tr\big(W_1\cdot H_\tau\cdot W_1\cdot H_\tau\big) - \sum_{t\in t_1}^{T_1} E\big[Tr\big(W_1\cdot\Delta y_t\cdot\Delta y_t'\cdot W_1'\cdot\Omega^{-1}\big)\big] = \\
&= -T_1\cdot Tr\big(W_1\cdot H_\tau\cdot W_1\cdot H_\tau\big) - \sum_{t\in t_1}^{T_1} Tr\big(W_1\cdot E\big[\Delta y_t\cdot\Delta y_t'\big]\cdot W_1'\cdot\Omega^{-1}\big) = \\
&= -T_1\cdot Tr\big(W_1\cdot H_\tau\cdot W_1\cdot H_\tau\big) - \sum_{t\in t_1}^{T_1} Tr\big(W_1\cdot E\big[H_\tau\cdot X_t\cdot\beta\cdot\varepsilon_t'\cdot H_\tau'+ \\
&+ H_\tau\cdot X_t\cdot\beta\cdot\beta'\cdot X_t'\cdot H_\tau' + H_\tau\cdot\varepsilon_t\cdot\varepsilon_t'\cdot H_\tau'\cdot\varepsilon_t\cdot\beta'\cdot X_t'\cdot H_\tau'\big]\cdot W_1'\cdot\Omega^{-1}\big) = \\
&= -T_1\cdot Tr\big(W_1\cdot H_\tau\cdot W_1\cdot H_\tau\big) - \sum_{t\in t_1}^{T_1} Tr\Big(W_1\cdot\Big[H_\tau\cdot X_t\cdot\beta\cdot E[\varepsilon_t']\cdot H_\tau'+ \\
&+ H_\tau\cdot X_t\cdot\beta\cdot\beta'\cdot X_t'\cdot H_\tau' + H_\tau\cdot E[\varepsilon_t\cdot\varepsilon_t']\cdot H_\tau' + E[\varepsilon_t]\cdot\beta'\cdot X_t'\cdot H_\tau'\Big]\cdot W_1'\cdot\Omega^{-1}\Big) = \\
&= -T_1\cdot Tr\big(W_1\cdot H_\tau\cdot W_1\cdot H_\tau\big) - \\
&- \sum_{t\in t_1}^{T_1} Tr\Big(W_1\cdot\Big[H_\tau\cdot X_t\cdot\beta\cdot\beta'\cdot X_t'\cdot H_\tau' + H_\tau\cdot\Omega\cdot H_\tau'\Big]\cdot W_1'\cdot\Omega^{-1}\Big) = \\
&= -T_1\cdot Tr\big(W_1\cdot H_\tau\cdot W_1\cdot H_\tau\big) - \\
&- \sum_{t\in t_1}^{T_1} Tr\big(W_1\cdot H_\tau\cdot X_t\cdot\beta\cdot\beta'\cdot X_t'\cdot H_\tau'\cdot W_1'\cdot\Omega^{-1} + W_1\cdot H_\tau\cdot\Omega\cdot H_\tau'\cdot W_1'\cdot\Omega^{-1}\big) = \\
&= -T_1\cdot Tr\big(W_1\cdot H_\tau\cdot W_1\cdot H_\tau + H_\tau'\cdot W_1'\cdot\Omega^{-1}\cdot W_1\cdot H_\tau\cdot\Omega\big) - \\
&- \sum_{t\in t_1}^{T_1} Tr\big(\beta'\cdot X_t'\cdot H_\tau'\cdot W_1'\cdot\Omega^{-1}\cdot W_1\cdot H_\tau\cdot X_t\cdot\beta\big) =
\end{aligned}
$$

Setting $M_1^\tau = H'_\tau \cdot W'_1 \cdot \Omega^{-1} \cdot W_1 \cdot H_\tau$ we can rewrite the above identity as:

$$\begin{aligned}
E[\mathscr{H}_{1,1}] &= -T_1 \cdot Tr\big(W_1 \cdot H_\tau \cdot W_1 \cdot H_\tau + M_1^\tau \cdot \Omega\big) - \sum_{t \in t_1}^{T_1} \beta' \cdot X'_t \cdot M_1^\tau \cdot X_t \cdot \beta = \\
&= -T_1 \cdot Tr\big(W_1 \cdot H_\tau \cdot W_1 \cdot H_\tau + M_1^\tau \cdot \Omega\big) - \beta' \cdot X'_\tau \cdot \big(I_{T_1} \otimes M_1^\tau\big) \cdot X_\tau \cdot \beta.
\end{aligned}$$

Simmetrically:

$$E[\mathscr{H}_{2,2}] = -T_2 \cdot Tr\big(W_2 \cdot H_\gamma \cdot W_2 \cdot H_\gamma + M_1^\gamma \cdot \Omega\big) - \beta' \cdot X'_\gamma \cdot \big(I_{T_2} \otimes M_1^\gamma\big) \cdot X_\gamma \cdot \beta.$$

with $M_1^\gamma = H'_\gamma \cdot W'_2 \cdot \Omega^{-1} \cdot W_2 \cdot H_\gamma$.

Going on with the calculation:

$$\begin{aligned}
E[\mathscr{H}_{1,3:1,23}] &= E\Big[-\sum_{t \in t_1}^{T_1} X'_t \cdot \Omega^{-1} \cdot W_1 \cdot \Delta y_t\Big] = \\
&= -\sum_{t \in t_1}^{T_1} X'_t \cdot \Omega^{-1} \cdot W_1 \cdot E\Big[H_\tau \cdot X_t \cdot \beta + H_\tau \cdot \varepsilon_t\Big] = \\
&= -\sum_{t \in t_1}^{T_1} X'_t \cdot \Omega^{-1} \cdot W_1 \cdot H_\tau \cdot X_t \cdot \beta \\
&= X'_\tau \cdot (I_{T_1} \otimes M_2^\tau) \cdot X_\tau \cdot \beta
\end{aligned}$$

with $M_2^\tau = \Omega^{-1} \cdot W_1 \cdot H_\tau$.

Simmetrically:

$$E[\mathscr{H}_{2,3:2,23}] = X'_\gamma \cdot (I_{T_2} \otimes M_2^\gamma) \cdot X_\gamma \cdot \beta$$

with $M_2^\gamma = \Omega^{-1} \cdot W_2 \cdot H_\gamma$.

Next step:

$$
\begin{aligned}
E[\mathscr{H}_{1,24:1,38}] &= Tr\Big( \big( \sum_{t\in t_1}^{T_1} E\big[\Delta y_t \cdot \varepsilon_t'\big]\big) \cdot \frac{\partial \Omega^{-1}}{\partial \sigma_i^2} \cdot W_1 \Big) = \\
&= Tr\Big( \big( \sum_{t\in t_1}^{T_1} E\big[\Delta y_t \cdot \varepsilon_t'\big]\big) \cdot \frac{\partial \Omega^{-1}}{\partial \sigma_i^2} \cdot W_1 \Big) = \\
&= Tr\Big( \big( \sum_{t\in t_1}^{T_1} H_\tau \cdot E\big[\varepsilon_t \cdot \varepsilon_t'\big]\big) \cdot \frac{\partial \Omega^{-1}}{\partial \sigma_i^2} \cdot W_1 \Big) = \\
&= T_1 \cdot Tr\Big( H_\tau \cdot \Omega \cdot \frac{\partial \Omega^{-1}}{\partial \sigma_i^2} \cdot W_1 \Big) = \\
&= T_1 \cdot Tr\Big( \Omega \cdot \frac{\partial \Omega^{-1}}{\partial \sigma_i^2} \cdot W_1 \cdot H_\tau \Big),
\end{aligned}
$$

Notice that

$$
\Omega \cdot \frac{\partial \Omega^{-1}}{\partial \sigma_i^{-2}} = -\sigma_i^2 \cdot I_{ii}
$$

where the generic element of matrix $I_{ii}$ is given by

$$
\omega_{s,t} = \begin{cases} 1 & s = i, j = i \\ 0 & \text{otherwise} \end{cases}
$$

Therefore

$$
\begin{aligned}
E[\mathscr{H}_{1,23+i}] &= T_1 \cdot \sigma_i^{-2} \cdot Tr\Big( I_{ii} \cdot W_1 \cdot H_\tau \Big) = \\
&= T_1 \cdot \sigma_i^{-2} \cdot \Big( W_1 \cdot H_\tau \Big)_{ii}
\end{aligned}
$$

Finally we have that:

$$
E[\mathscr{H}_{1,24:1:38}] = T_1 \cdot diag\Big( \Omega^{-1} \cdot W_1 \cdot H_\tau \Big) = T_1 \cdot diag(M_2^\tau).
$$

Simmetrically:

$$
E[\mathscr{H}_{2,24:2:38}] = T_2 \cdot diag\Big( \Omega^{-1} \cdot W_2 \cdot H_\gamma \Big) = T_2 \cdot diag(M_2^\gamma).
$$

Going on:

$$E[\mathscr{H}_{3,3:23,23}] = E[\sum_{t=1}^{T} X_t' \cdot \Omega^{-1} \cdot X_t] = \sum_{t=1}^{T} X_t' \cdot \Omega^{-1} \cdot X_t = X' \cdot \Sigma^{-1} \cdot X$$

$$\begin{aligned} E[\mathscr{H}_{3,24:23,38}] &= E[\sum_{t=1}^{T} X_t' \cdot \frac{\partial \Omega^{-1}}{\partial \sigma^2} \cdot \varepsilon_t] \\ &= \sum_{t=1}^{T} X_t' \cdot \frac{\partial \Omega^{-1}}{\partial \sigma^2} \cdot E[\varepsilon_t] \\ &= \underset{k \times n}{\mathbf{0}} \end{aligned}$$

$$E[\mathscr{H}_{23+i,23+j|i,j\in[1,n]}] = \begin{cases} \frac{T}{2} \cdot \frac{1}{\sigma_i^4} \cdot \left(1 - \frac{2}{T \cdot \sigma_i^2} \cdot \sum_{t=1}^{T} E[\varepsilon_{i,t}^2]\right) & \forall i = j \in [1,n] \\ \\ 0 \quad \forall i \neq j \end{cases}$$

$$\begin{aligned} &= \begin{cases} -\frac{T}{2} \cdot \frac{1}{\sigma_i^4} & \forall i = j \in [1,n] \\ 0 \quad \forall i \neq j \end{cases} \\ &= -\frac{T}{2} \cdot \begin{bmatrix} \sigma_1^{-4} & 0 & \cdots & 0 \\ 0 & \sigma_2^{-4} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ o & 0 & \cdots & \sigma_n^{-4} \end{bmatrix} = -\frac{T}{2} \cdot V \end{aligned}$$

We finally have all the elements of the Fisher Information Matrix for our panel (with dummy variables) spatial model:

$$\mathscr{I} =$$

$$
\begin{bmatrix}
\begin{array}{c}-T_1\cdot Tr\big(W_1\cdot H_\tau\cdot W_1\cdot H_\tau+M_1^\tau\cdot\Omega\big)-\\ -\beta'\cdot X_\tau'\cdot\big(I_{T_1}\otimes M_1^\tau\big)\cdot X_\tau\cdot\beta\end{array} & 0 & \big(X_\tau'\cdot(I_{T_1}\otimes M_2^\tau)\cdot X_\tau\cdot\beta\big)' & T_1\cdot diag(M_2^\tau)' \\
0 & \begin{array}{c}-T_2\cdot Tr\big(W_2\cdot H_\gamma\cdot W_2\cdot H_\gamma+M_1^\gamma\cdot\Omega\big)-\\ -\beta'\cdot X_\gamma'\cdot\big(I_{T_2}\otimes M_1^\gamma\big)\cdot X_\gamma\cdot\beta\end{array} & \big(X_\gamma'\cdot(I_{T_2}\otimes M_2^\gamma)\cdot X_\gamma\cdot\beta\big)' & T_2\cdot diag(M_2^\gamma)' \\
X_\tau'\cdot(I_{T_1}\otimes M_2^\tau)\cdot X_\tau\cdot\beta & X_\gamma'\cdot(I_{T_2}\otimes M_2^\gamma)\cdot X_\gamma\cdot\beta & X'\cdot\Sigma^{-1}\cdot X & \underset{k\times n}{\mathbf{0}} \\
T_1\cdot diag(M_2^\tau) & T_2\cdot diag(M_2^\gamma) & \underset{n\times k}{\mathbf{0}} & -\dfrac{T}{2}\cdot V
\end{bmatrix}
$$

## B.3 Model Selection - Vuong Test for Static Spatial Panel Data

In this section we explain how we implement the Vuong Test for model selection. Firstly, the Vuong test (see Vuong (1989)) is meant to discriminate between two misspecified and non-nested models. Basically, we assume there is a hidden true model and we want to choose one of two competing non-nested models which fit the data equally well. The Vuong test calculates and compares the Kullback-Leibler distance between the two and the true model. In practice, is a t-test on the KL divergence. One problem we encounter is that it was developed for one-dimensional iid data, however, we deal with a panel whose observations are serially uncorrelated but spatially correlated. Wooldridge (2010) shows that the Vuong test can easily be extended to panel data models by accounting for serial correlation in the time series.[30] However, in our problem the $n \times 1$ vector of industry observations is iid over time and our asymptotic keeps the cross-sectional dimension, which is spatially correlated, fixed, and then let the time series to go to infinite $T \to \infty$. Economically speaking this makes sense: we observe those fixed 62 industries over time, however, the cross sectional dimension exceeds the times series one, 37 years. This means that our finite sample distribution will not be a very good approximation of the asymptotic one. However, this is the best we can do, given the data availability.

Let's derive now the Vuong Test. The quasi-log-likelihood of the baseline model, Equation (2), is:

$$\ell_{t,B}(\underbrace{\rho, \beta, \Omega}_{\theta_B}) = \log f_B(\Delta y_t | X_t; \theta_B) = -\frac{n}{2} \ln(2\pi) - \frac{1}{2} \cdot \ln(|\Omega|) + \\ + \ln(|I_n - \rho^{down} \cdot A \cdot TB_t - \rho^{up} \cdot \hat{A}' \cdot EB_t|) - \frac{1}{2} \cdot \varepsilon_t' \cdot \Omega^{-1} \cdot \varepsilon_t.$$

with:

$$\varepsilon_t = \Big( I_n - \rho^{down} \; A \; TB_t - \rho^{up} \; \hat{A}' \; EB_t \Big) \cdot \Delta y_t - X_t \cdot \beta.$$

The sum of the quasi-log-likelihood evaluated at the MLE, $\hat{\theta}_B$, for the baseline model is: $\mathscr{L}_B = \sum_{t=1}^{T} \ell_{t,B}(\hat{\theta}_B)$. Analogously, for the inverted model, Equation

[30] See Section 13.11.2 - Model Selection Tests.

(3), the quasi-log-likelihood is:

$$\ell_{t,I}(\underbrace{\tilde{\rho},\tilde{\beta},\tilde{\Omega}}_{\theta_I}) = \log f_I(\Delta y_t|X_t;\theta_I) = -\frac{n}{2}\ln(2\pi) - \frac{1}{2}\cdot\ln(|\tilde{\Omega}|) +$$
$$+\ln(|I_n - \tilde{\rho}^{down}\cdot A\cdot EB_t - \tilde{\rho}^{up}\cdot \hat{A}'\cdot TB_t|) - \frac{1}{2}\cdot\varepsilon_t'\cdot\tilde{\Omega}^{-1}\cdot\varepsilon_t.$$

with:

$$\varepsilon_t = \Big(I_n - \tilde{\rho}^{down}\ A\ EB_t - \tilde{\rho}^{up}\ \hat{A}'\ TB_t\Big)\cdot\Delta y_t - X_t\cdot\tilde{\beta}.$$

The sum of the quasi-log-likelihood evaluated at the MLE, $\hat{\theta}_I$, for the inverted model is: $\mathscr{L}_I = \sum_{t=1}^T \ell_{t,I}(\hat{\theta}_I)$.
Following Wooldridge (2010), let's define the estimator for the variance of the KL divergence as:

$$\hat{\eta}^2 = \frac{1}{T}\cdot\sum_{t=1}^{T}\Big(\ell_{t,B}(\hat{\theta}_B) - \ell_{t,I}(\hat{\theta}_I)\Big)^2.$$

Then, the Vuong Model Selection Statistic, VMS, is:

$$\begin{aligned} VMS &= T^{-1/2}\cdot\frac{(\mathscr{L}_B - \mathscr{L}_I)}{\hat{\eta}} \\ &= \frac{\frac{1}{T}\cdot\sum_{t=1}^T\Big(\ell_{t,B}(\hat{\theta}_B) - \ell_{t,I}(\hat{\theta}_I)\Big)}{\sqrt{\dfrac{\frac{1}{T}\cdot\sum_{t=1}^T\Big(\ell_{t,B}(\hat{\theta}_B) - \ell_{t,I}(\hat{\theta}_I)\Big)^2}{T}}} \xrightarrow{d} N(0,1) \end{aligned}$$

where the standard normal distribution holds under:

$$H_0 : \mathbb{E}[\ell_{t,B}(\theta_B^*)] = \mathbb{E}[\ell_{t,I}(\theta_I^*)]$$

where $\theta_B^*$ and $\theta_I^*$ are the pseudo-true values of the parameters. Basically, the null hypothesis is saying that the two potentially misspecified models fit the data equally well. Notice that the test is super easy to implement: 1) define the difference: $\hat{d}_t = \ell_{t,B}(\hat{\theta}_B) - \ell_{t,I}(\hat{\theta}_I)$; 2. Regress $\hat{d}_t$ on unity; 3. Run a t-test to verify that the average of the difference is statistically different from zero. We reject the null hypothesis in favor of a better fit to the data of the baseline model if $\hat{d}_t$ is statistically greater than zero. Notice that if this happens it does not mean that the baseline model is correctly specified (although it could be), however, we can conclude that the baseline model fits better in terms of

expected likelihood.
The value we obtain is VMS = 0.033 which is clearly not statistically different from zero. Even if positive sign of the statistics points at a better fit of the baseline model against the inverted one, there is not enough statistical evidence to claim that the baseline outperforms on average the inverted model.

## B.4 Bayesian MCMC - Technical Details

Even if the MLE is a common standard method in spatial econometric applications, we have two valid reasons for not adopting it: 1. non-stationary estimates of aggregate total effects; 2. prior information on the values of the parameters. Let's explore both the issues.

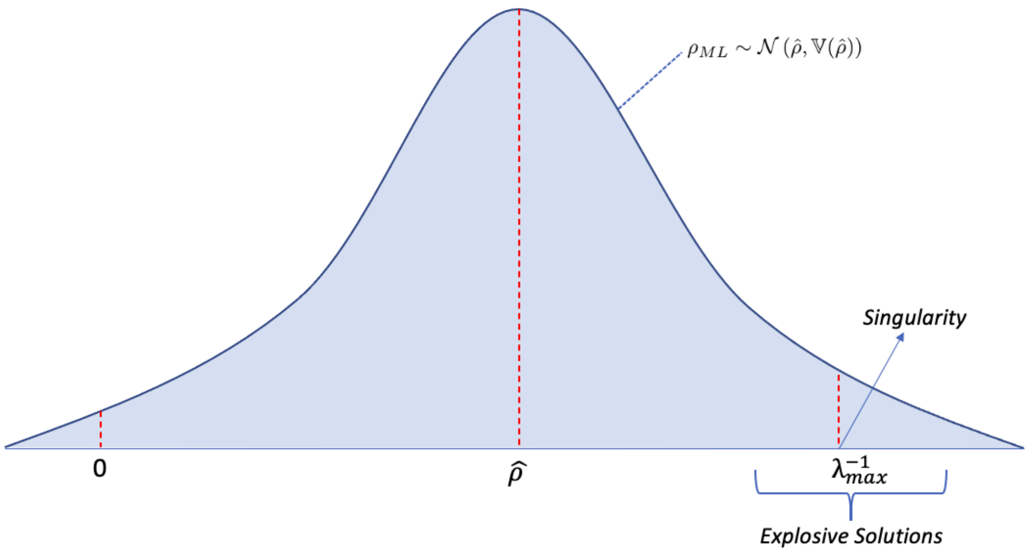
### 1. Non-Stationary Solutions

We can estimate the parameters by maximizing the concentrated log-likelihood over the compact set which guarantees a positive definite matrix (see Ord (1975)): $C^{down} = \left(\lambda_{\min}^{-1}, \lambda_{\max}^{-1}\right)$ and $C^{up} = \left(\mu_{\min}^{-1}, \mu_{\max}^{-1}\right)$. The standard errors are constructed using the analytical Fisher Information of the model, centered on the point estimates, $\hat{\rho}^{down}$ and $\hat{\rho}^{up}$. The asymptotic results of Yu, De-Jong, and Lee (2008) guarantees the asymptotic normality of the parameters of equation (2) and (3) (See Theorem 3 case $n/T \to 0$). For instance, for the estimator of $\rho^{down}$ we have:

$$\sqrt{T \cdot n}\left(\hat{\rho}_{nT}^{down} - \rho^{down}\right) \xrightarrow{d} \mathscr{N}\left(0, \sigma^2\right)$$

where $\sigma^2$ is the asymptotic variance of the MLE, obtained by the calculating the analytical Fisher Information matrix of our model. However, we are interested in estimating the aggregate total effect of fiscal consolidations, not the parameters of the model themselves. At page 70, LeSage and Pace (2009) suggest to construct the asymptotic distribution of the average total effect (our aggregate total effect) by following these steps: 1. estimate the parameters of the model via MLE; 2. Draw values of the parameters by their approximate asymptotic distribution ($\tilde{\rho}^{down} \approx \mathscr{N}\left(\hat{\rho}_{nT}^{down}, \frac{\hat{\sigma}^2(\hat{\rho}_{nT}^{down})}{nT}\right)$; 3. Calculate at each step the aggregate total effect. After doing so we calculated the standard errors of the $ATE_{TB}$ by calculating the standard deviation of the asymptotic distribution so constructed. We obtained explosive solutions. This is a surprising result, in fact, the asymptotic normality of the average effect is guaranteed by

Figure 6: Explosive Solutions of $ATE_{TB}$




the $\Delta$-method:

$$\sqrt{T \cdot n}\left(ATE_{TB}(\hat{\rho}^{down}_{nT}) - ATE_{TB}(\rho^{down})\right) \xrightarrow{d} \mathcal{N}\left(0, \sigma^2 \cdot \left(\frac{\partial ATE_{TB}(\rho^{down})}{\partial \rho^{down}}\right)^2\right)$$

where $ATE_{TB} : C^{down} \rightarrow \mathbb{R}$ and $ATE_{TB}(x) = v' \cdot (I_n - x \cdot A)^{-1} \cdot \omega_{TB}$ and $v$ is a vector of industry output shares of total industrial production (the weights we use to calculate the aggregate effect of fiscal consolidations). What goes wrong in this procedure? The $\Delta$-method is an asymptotic result, which might provide a terrible approximation of a finite sample distribution. It all boils down in finding a distribution which approximates well the small sample one. If $\hat{\rho}^{down}_{nT}$ is very closed to the boundary and its asymptotically normal standard errors are large, that is, they approach the boundary of $C^{down}$ then we end up drawing values of $\rho^{down}$ which deliver unrealistically large values of $ATE_{TB}$, because matrix $(I_n - \rho^{down} \cdot A)^{-1}$ becomes singular (the boundary is one eigenvalue of $A$). This situation is described in Figure 6.

## 2. Prior Information

We have two extra "prior" pieces of information on the value of the spatial parameters, $\rho^{down}$ and $\rho^{up}$:

i. Values of $\rho^{down}$ and $\rho^{up}$ close to the boundaries will deliver unrealistically high values of ATE, ADE and ANE, since the determinant of matrices

$(I_n - \rho^{down} \cdot A)$ and $(I_n - \rho^{up} \cdot \hat{A}^T)$ will approach zero by definition of eigenvalue. In turn, the elements of their inverse matrices will explode, as illustrated above. Therefore, we should assign less weight to values of $\rho^{down}$ and $\rho^{up}$ close to the boundaries.

ii. We know that industries that are close to each other in the production network will co-move. For instance, if industry X faces increasing prices for its input, it will shrink production and increase prices; in turn, customers of X will also face the same problem and will react similarly, by reducing production and increasing prices. Therefore, the direction of the spatial correlation among industries' output is positive: $\rho^{down} > 0$ and $\rho^{up} > 0$.

### Model Estimation

We can integrate such prior information into our estimation and avoid non-stationarity aggregate effects, by adopting a Bayesian MCMC similar to the one introduced by LeSage and Parent (2007). We illustrate here how we implement the Bayesian MCMC to estimate the parameters of Equation (2) (baseline). The log-likelihood of that model is the one outlined above. The priors we employ on the parameters are:

$$
\begin{aligned}
&\pi(\beta) \propto constant \\
&\Omega = \sigma^2 \cdot V \quad \text{with } V = diag(v_1, ..., v_n) \\
&\pi(\sigma^2) \propto \frac{1}{\sigma^2} \\
&\pi(v_i) \overset{iid}{\sim} \Gamma^{-1}\left(\frac{r}{2}, \frac{r}{2}\right), \quad i = 1, ..., n \\
&\rho^{down} \sim Gen.Beta(d, d) \\
&\rho^{up} \sim Gen.Beta(d, d).
\end{aligned}
$$

We adopt non-informative priors for $\sigma^2$ and $\beta$ to reflect our lack of information around the values of these parameters. Concerning $r$, a lower value generates more diffusion in the distributions of $v_i$, thus regulating our confidence towards heteroskedasticity. Unlike LeSage and Pace (2009), who suggest a value of 4, we set $r$ equal to 3 to reflect a strong belief towards heteroskedasticity. For instance, industries in the Agriculture (NAICS 11) as well as Mining (NAICS 21) macro sectors, exhibit much higher volatilities than the rest of the industries.

We impose a "generalized (or non-standardized) $Beta(d, d)$ prior", with support from 0 to $\lambda_{max}^{-1}$ for $\rho^{down}$ and from 0 to $\hat{\lambda}_{max}^{-1}$ for $\rho^{up}$. We follow LeSage and Pace (2009) and set $d$ equal to 1.1; which has the benefit of letting the generalized Beta prior to resemble a Uniform distribution (diffuse prior), but with low density at the boundaries, as illustrated in Figure 7. The choice of

Figure 7: Generalized Beta prior

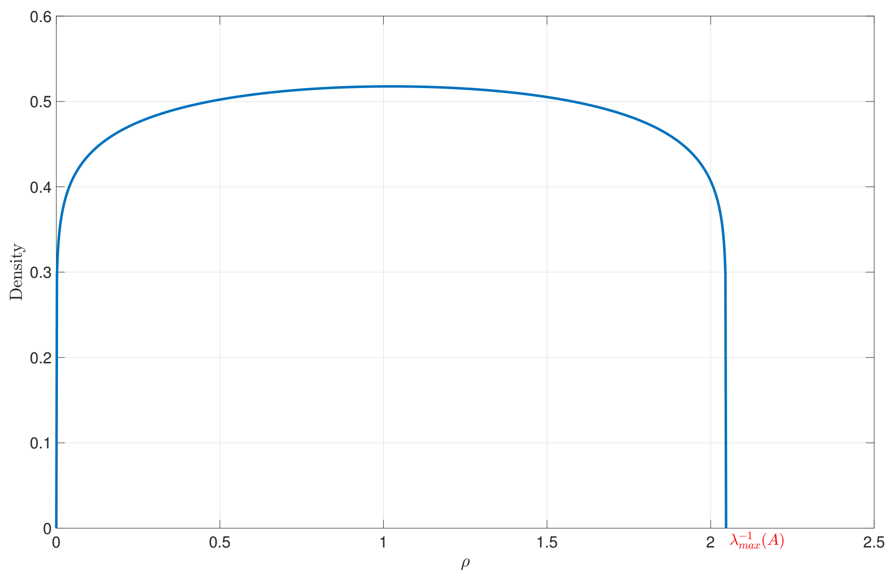



*Figure 7: line-plot of a non-standardized $Beta(1.1, 1.1)$ density function, with support from $(0, \lambda_{max}^{-1}(A) = 2.047)$ which we employ as a prior for the spatial parameter $\rho^{down}$.*

such a prior allows us to be agnostic about the specific value of the spatial parameters but at the same time it allows to embed the prior information we have into their estimates.

Furthermore, we assume that all the prior distributions are independent from each other. We use the standard "Metropolis within Gibbs" algorithm, and we obtain an approximation of the posterior densities for each parameter of the model.

We now outline the precise steps of the procedure:

1. **Initialization:** Set up initial values for the parameters: $\beta_{(0)}, \sigma^2_{(0)}, V_{(0)}, \rho^{down}_{(0)}, \rho^{up}_{(0)}$, where $V_{(0)} = diag\left(v^2_{1,(0)}, ..., v^2_{n,(0)}\right)$.

2. **Gibbs Sampling:**

a) Draw $\beta_{(1)}$ from the conditional posterior distribution, which is obtained by mixing the likelihood with a normal prior with mean $c$ (a vector of zeros in our simulation) and covariance matrix $L$. In order to not add any information, we simply set $L$ to be equal to a diagonal matrix whose entries are infinite (1e12 in our simulation):

$$P(\beta_{(0)}|\mathscr{D}, \sigma^2_{(0)}, V_{(0)}, \rho^{down}_{(0)}, \rho^{up}_{(0)}) = \mathscr{N}(c^*, L^*) \propto \mathscr{L}(\theta|\mathscr{D}) \cdot \mathscr{N}(c, L)$$

$$c^* = \frac{1}{T} \cdot (\sum_{t=1}^{T} X'_t \cdot V^{-1}_{(0)} \cdot X_t + \frac{\sigma^2_{(0)}}{T} \cdot L^{-1})^{-1} \cdot (\frac{1}{T} \cdot \sum_{t=1}^{T} X'_t \cdot V^{-1}_{(0)} \cdot H_t \cdot \Delta y_t + \frac{\sigma^2_{(0)}}{T} \cdot L^{-1} \cdot c)$$

$$L^* = \frac{\sigma^2_{(0)}}{T} \cdot (\sum_{t=1}^{T} X'_t \cdot V^{-1}_{(0)} \cdot X_t + \frac{\sigma^2_{(0)}}{T} \cdot L^{-1})^{-1}$$

b) Draw $\sigma^2_{(1)}$ from the conditional posterior distribution, which is proportional to likelihood times an inverse gamma distribution as a prior:

$$P(\sigma^2_{(1)}|\mathscr{D}, \beta_{(1)}, V_{(0)}, \rho^{down}_{(0)}, \rho^{up}_{(0)}) = \Gamma^{-1}(\frac{\theta_1}{2}, \frac{\theta_2}{2}) \propto \mathscr{L}(\theta|\mathscr{D}) \cdot \Gamma^{-1}(a, b)$$

$$\theta_1 = nT + 2a \quad \theta_2 = \sum_{t=1}^{T} \varepsilon'_t \cdot V^{-1}_{(0)} \cdot \varepsilon_t + 2b$$

In practice we draw $\sigma^2_{(1)}$ from $\theta_2/\chi_{\theta_1}$.
Notice that, setting $a$ and $b$ (the prior's parameters) equal to 0, is like putting a Jefferey's prior on $\sigma^2$. This is exactly what we do.

c) Draw $v_{i,(1)}$ from the following conditional posterior distribution, proportional to an inverse gamma prior:

$$P(v_{i,(1)}|\mathscr{D}, \sigma^2_{(1)}, \rho^{down}_{(0)}, \rho^{up}_{(0)}) = \Gamma^{-1}(\frac{q_1}{2}, \frac{q_2}{2}) \propto \mathscr{L}(\theta|\mathscr{D}) \cdot \Gamma^{-1}(\frac{r}{2}, \frac{r}{2})$$

$$q_1 = r + T \quad q_2 = \frac{1}{\sigma^2_{(1)}} \cdot \sum_{t=1}^{T} \varepsilon^2_{i,t} + r$$

In practice we draw $v_{i,(1)}$ from $q_2/\chi_{q_1}$.
As anticipated above in the paper, since we are confident on the heteroskedastic behavior of industry value added, we set our prior hyperparameter $r$ to be equal to 3 rather than 4, as done in LeSage and Pace (2009).
Replicating this procedure $n$ times, we get a first simulation of matrix $V_{(1)}$.

3. **Metropolis-Hastings**: We now need to draw the spatial coefficients. However we cannot apply a simple Gibbs Sampling, since the conditional posterior distribution is not defined for them. LeSage and Pace (2009) suggest the adoption of the Metropolis-Hastings algorithm to overcome this problem. To ease notation we set $\rho_1 := \rho^{down}$ and $\rho_2 := \rho^{up}$. The algorithm is the following:

   (a) Draw $\rho_1^c$ (where the $c$ superscript stands for "candidate") from the (random walk) proposal distribution:

   $$\rho_1^c = \rho_{1,(0)} + c_1 \cdot \mathscr{N}(0,1)$$

   (b) Run a bernoulli experiment to determine the updated value of $\rho_1$:

   $$\rho_{1,(1)} = \begin{cases} \rho_1^c & \pi \quad \text{(accept)} \\ \rho_{1,(0)} & 1-\pi \quad \text{(reject)} \end{cases}$$

   Where $\pi$ is equal to $\pi = \min\{1, \psi_{MH_1}\}$ and, setting: $A_\tau(\rho_1) = I_n - \rho_1 \cdot W_1$, we have:

   $$\begin{aligned}\psi_{MH_1} = \frac{|A_\tau(\rho_1^c)|}{|A_\tau(\rho_{1,(0)})|} \cdot \exp\Big\{ &- \frac{1}{2\sigma^2_{(1)}} \cdot \sum_{t \in t_1}^{T_1} \Big[ \Delta y_t' \cdot \big( A_\tau(\rho_1^c)' \cdot V_{(1)}^{-1} \cdot A_\tau(\rho_1^c) - \\ &- A_\tau(\rho_{1,(0)})' \cdot V_{(1)}^{-1} \cdot A_\tau(\rho_{1,(0)}) \big) \cdot \Delta y_t - \\ &- 2\beta' \cdot X_t' \cdot V_{(1)}^{-1} \big( A_\tau(\rho_1^c) - A_\tau(\rho_{1,(0)}) \big) \cdot \Delta y_t \Big] \Big\} \cdot \\ &\cdot \left[ \frac{(\rho_1^c - 0) \cdot (\lambda_{max}^{-1} - \rho_1^c)}{(\rho_{1,(0)} - 0) \cdot (\lambda_{max}^{-1} - \rho_{1,(0)})} \right]^{d-1} \cdot \mathbf{1}\big(0 \le \rho_1^c \le \lambda_{max}^{-1}\big)\end{aligned}$$

   Basically, we compute the probability to accept the candidate value from the proposal distribution, and then we update the value of $\rho_1$ by running the bernoulli experiment with such a probability of success. Notice that if we draw a value of $\rho_1$ outside the support of the beta prior, $\psi_{MH_1} = 0$ and then $\pi = 0$ and we clearly reject the candidate value.
   We set $d$ equal to 1.1, on both $\rho_1$ and $\rho_2$; this is done to resemble a Uniform (0,1) but with less density on its boundary values.

   (c) Once updated $\rho_1$, we replicate the procedure for $\rho_2$. Setting $A_\gamma(\rho_2) =$

$I_n - \rho_2 \cdot W_2$ we have:

$$\psi_{MH_2} = \frac{|A_\gamma(\rho_2^c)|}{|A_\gamma(\rho_{2,(0)})|} \cdot \exp\Big\{ - \frac{1}{2\sigma^2_{(1)}} \cdot \sum_{t \in t_2}^{T_2} \Big[ \Delta y_t' \cdot \big( A_\gamma(\rho_2^c)' \cdot V_{(1)}^{-1} \cdot A_\gamma(\rho_2^c) -$$
$$- A_\gamma(\rho_{2,(0)})' \cdot V_{(1)}^{-1} \cdot A_\gamma(\rho_{2,(0)}) \big) \cdot \Delta y_t -$$
$$- 2\beta' \cdot X_t' \cdot V_{(1)}^{-1} \big( A_\gamma(\rho_2^c) - A_\gamma(\rho_{2,(0)}) \big) \cdot \Delta y_t \Big] \Big\} \cdot$$
$$\cdot \left[ \frac{(\rho_2^c - 0) \cdot (\hat{\lambda}_{max}^{-1} - \rho_2^c)}{(\rho_{2,(0)} - 0) \cdot (\hat{\lambda}_{max}^{-1} - \rho_{2,(0)})} \right]^{d-1} \cdot \mathbf{1}\big(0 \leq \rho_2^c \leq \hat{\lambda}_{max}^{-1}\big)$$

(d) At this point we need to update the variance of the proposal distributions: if the acceptance rate (number of acceptances over number of iterations of the Markov Chain) of the first parameter $\rho_1$ falls below 40% we need to reduce the value of $c_1$, the so called tuning parameter, which regulates the variance of the proposal distribution. The variance is reduced by rescaling it: $c_1' = \frac{c_1}{1.1}$. In this way, we are able to draw values closer to the current state of $\rho_1$, and therefore, we expect to increase the acceptance rate.
On the contrary, if the acceptance rate rises above 60%, we need to increase the tuning parameter, in order to draw values far from the current state, in this way we increase the chance to explore more the low-density parts of the distribution. We increase the variance of the candidate distribution by scaling upward its standar deviation: $c_1' = 1.1 \cdot c_1$.
Clearly we replicate this procedure also for $\rho_2$.

4. **Repeat**: Once updated all the values, we replicate steps 2 and 3, 45,000 times to make sure the acceptance rate has converged.

5. **Burn-in**: we drop the first 35,000 iterations of the Markov Chain, thus obtaining a vector of 10,000 observations for each of the parameters, which account for the simulated posterior distributions.

## B.5 Simulating the ATE, ADE and ANE

We construct via Monte Carlo the distribution of the ATE, ADE and ANE. In particular we follow these steps:

1. **(Parameters)** Draw $\rho^{down}$, $\rho^{up}$, $\boldsymbol{\tau}$ and $\boldsymbol{\gamma}$ from their posterior distributions. To take into account the potential correlation among them, draw from the same iteration of the Bayesian MCMC.

2. **(Style of the plan)** Construct both a TB and an EB simulated fiscal plan, by drawing the style from a distribution which mimics the empirical one.

3. **(Average effects)** Construct ATE, ADE and ANE using the parameters drawn in step 1 and the style drawn in step 2.

4. Repeat 100,000 times steps from 1 though 3, to make sure all the possible combination of styles and parameters are simulated.

Step 2 allows us to claim that the baseline results reported in the paper are robust to different styles of fiscal plans.

## Empirical distribution of style of fiscal plans

We are interested in simulating a 2 years fiscal consolidation made of an unexpected part, no announced part and a single year future part to be implemented in the second year of the simulation.
First of all, we want to simulate the unexpected part of the fiscal plan, therefore, we need to look at those years when an unanticipated shock occurs. Define the two sub-samples: $TB^u := \{t : 1, ..., T \mid tax_t^u > 0\}$ and $EB^u := \{t : 1, ..., T \mid exp_t^u > 0\}$. Then calculate the mean and the standard deviation of the unexpected component conditional on the occurrence of an unexpected shock:

$$\mu_\tau := \mathbb{E}(tax_t^u \mid t \in TB^u) \qquad \sigma_\tau := \sqrt{\mathbb{V}(tax_t^u \mid t \in TB^u)}$$
$$\mu_\gamma := \mathbb{E}(exp_t^u \mid t \in EB^u) \qquad \sigma_\gamma := \sqrt{\mathbb{V}(exp_t^u \mid t \in EB^u)}$$

In order to simulate a plausible unexpected component of the plan, we draw them from the following distributions:

$$\tilde{tax}^u \sim \mathscr{U}(\mu_\tau - \sigma_\tau, \mu_\tau + \sigma_\tau)$$
$$\tilde{exp}^u \sim \mathscr{U}(\mu_\gamma - \sigma_\gamma, \mu_\gamma + \sigma_\gamma)$$

where the $\tilde{}$ denotes a simulated component.
Concerning the future component, we need to predict what is the value of a one year ahead policy change, conditional on the occurrence of an unexpected policy change. Therefore, we run the following regressions:

$$tax_{t,1}^f = a_\tau + b_\tau \cdot tax_t^u \qquad \textit{with: } t \in TB^u$$
$$exp_{t,1}^f = a_\gamma + b_\gamma \cdot exp_t^u \qquad \textit{with: } t \in EB^u$$

The estimates of $a_\tau$, $b_\tau$, $a_\gamma$, $b_\gamma$ will be stored and used to predict values of $tax^f_{t,1}$ and $exp^f_{t,1}$, conditional on the occurrence of an unexpected component. At this point we have all the ingredients to outline the steps we do in the construction of a simulated style of the plan:

1. Draw unexpected components from their candidate distributions: $\tilde{tax}^u \sim \mathscr{U}(\mu_\tau - \sigma_\tau, \mu_\tau + \sigma_\tau)$ and $\tilde{exp}^u \sim \mathscr{U}(\mu_\gamma - \sigma_\gamma, \mu_\gamma + \sigma_\gamma)$.

2. Predict the future component using the estimates of $a_\tau$, $b_\tau$, $a_\gamma$, $b_\gamma$. We have: $\tilde{tax}^f = \hat{a}_\tau + \hat{b}_\tau \cdot \tilde{tax}^u$ and $\tilde{exp}^f = \hat{a}_\gamma + \hat{b}_\gamma \cdot \tilde{exp}^u$.

3. Normalize the value to one: $\tilde{tax}^u + \tilde{tax}^f = 1$ and $\tilde{exp}^u + \tilde{exp}^f = 1$.

For each iteration of the MC simulation used to approximate the posterior distributions of the ATE, ADE and ANE, we repeat steps 1 through 3 to simulate the style of the plan.
In the first year of the simulation we calculate the effects of TB and EB plans with style given by: $\boldsymbol{s}_{TB} = [\tilde{tax}^u \; 0 \; \tilde{tax}^f]$ and $\boldsymbol{s}_{EB} = [\tilde{exp}^u \; 0 \; \tilde{exp}^f]$ respectively. In the second year of the simulation, the future component of the shock is rolled over and becomes an announced and implemented shock. Therefore we calculate the effects of TB and EB plans with style given by: $\boldsymbol{s}_{TB} = [0 \; \tilde{tax}^f \; 0]$ and $\boldsymbol{s}_{EB} = [0 \; \tilde{exp}^f \; 0]$ respectively.

# C A Potential Theoretical Framework

We show here the theoretical framework which we have in mind when we refer to the theoretical transmission of demand and supply shocks. The model is a slight modification of Acemoglu, Akcigit, and Kerr (2016), which we adapted to allow for the propagation of a production tax.

The model considers a perfectly competitive economy with $n$ sectors, where the market clearing condition for the generic industry $i$ is:

$$y_i = c_i + \sum_{j=1}^{n} x_{ji} + G_i \tag{5}$$

where $c_i$ is household's consumption of good produced by industry $i$; $x_{ij}$[31] is the quantity of goods produced in industry $j$ used as inputs by industry

[31] In Equation (5) we actually have $x_{ji}$, that is, the amount of good $i$ used as input by industry $j$; we then sum over the $j$-s to obtain the total demand of good $i$ from all the industries.

$i$; $G_i$ are government purchases, funded by imposing either a lump sum or a sales-type tax:[32]

$$\sum_{i=1}^{n} p_i G_i = T + \tau \sum_{i=1}^{n} p_i y_i \tag{6}$$

Each sector solves the following profit maximization problem:

$$\max_{l_i, \{x_{ij}\}_{j=1}^{n}} (1-\tau) \cdot p_i \cdot y_i - w l_i - \sum_{j=1}^{n} p_j x_{ij}$$

with a production function similar to the one in Acemoglu, Carvalho, et al. (2012) and Carvalho (2014):[33]

$$y_i = l_i^{\alpha_i^l} \cdot (\prod_{j=1}^{n} x_{ij}^{\alpha_{ij}})^{\rho}$$

All alpha's are non negative, and we assume constant return to scale: $\alpha_i^l + \rho \cdot \sum_{j=1}^{n} a_{ij} = 1$. Notice here, that thanks to the Cobb-Douglas specification, $\rho$ can be interpreted as the share of intermediates in production.

The economy is populated by a representative agent, who maximizes utility subject to a budget constraint:

$$\max_{l, \{c_i\}_{i=1}^{n}} (1-l)^{\lambda} \cdot \prod_{i=1}^{n} c_i^{\beta_i}$$

$$\text{s.t.} \sum_{i=1}^{n} p_i c_i \leq wl - T$$

with $\sum_{i=1}^{n} \beta_i = 1$.
Firms and households take all prices as given, and the market-clearing conditions are satisfied in the goods market and the labor market. Government actions are taken as given and the wage is chosen as a numeraire $(w = 1)$.

[32]For example, an excise is a special type of sales tax, which is sector-specific. Excise tax might be of two types: ad valorem (percentage of values of a good) and specific (tax paid per unit). The excise tax may be paid by the producer, retailer, and consumer. Moreover, it might be taken on federal, state, and local levels.

[33]We omit the productivity component because we are not interested in studying productivity shocks.

## C.1 Network effect of a tax shock

By log-differentiating the equations which characterize the equilibrium the following closed-form expression of a tax shock effect is obtained:[34]

$$d\log y_i = d\log(1-\tau) + \alpha_i^l \cdot d\log(1-T) + \rho \quad \underbrace{\cdot \sum_{j=1}^{n} a_{ij} \cdot d\log y_j}_{\text{downstream spatial variable}} \tag{7}$$

Now, we introduce the input-output matrix A which collects all the coefficients of the Cobb-Douglas production function, in this way we can rewrite equation (7) in matrix form:

$$\underset{n\times 1}{d\log \mathbf{y}} = \mathbf{1}_n \cdot d\log(1-\tau) + \underset{n\times 1}{\boldsymbol{\alpha}^l} \cdot d\log(1-T) + \rho \cdot \underset{n\times n}{A} \cdot \underset{n\times 1}{d\log \mathbf{y}}, \tag{8}$$

where $\mathbf{1}_n$ denotes a column vector of ones of length $n$.
The above expression can be simplified by collecting the dependent variable $d\log \boldsymbol{y}$ on the left hand side of the expression. By doing this, we obtain the following closed form expression[35]:

$$d\log \boldsymbol{y} = -\underbrace{(I_n - \rho \cdot A)^{-1} \cdot \mathbf{1}_n}_{\text{downstream propagation}} \cdot \frac{\tau}{1-\tau} \cdot d\log\tau - \mathbf{1}_n \frac{T}{1-T} \cdot d\log T \tag{9}$$

The sectoral propagation of a tax adjustment is driven by the elements in the rows of the matrix $H := (I_n - \rho \cdot A)^{-1}$, which represents the Leontief inverse matrix. Notice that $H \cdot \mathbf{1}_n = \mathbf{1}_n + \rho \cdot A \cdot \mathbf{1}_n + \rho^2 \cdot A^2 \cdot \mathbf{1}_n + ...$, therefore, the downstream propagation depends on the rows of $A$, and describe how much intermediates sector $i$ purchases from all other sectors. We can see this from the FOC of firm $i$ with respect to $x_{ij}$:

$$a_{ij} \propto \frac{p_j \cdot x_{ij}}{p_i \cdot y_i} \approx \frac{\text{SALES}_{j\to i}}{\text{SALES}_i} \tag{10}$$

Therefore, the network propagation mechanism of a sales-type tax shock propagates downstream: at first each sector is hit by a tax shock; then firms re-optimize and increase their own price; by consequence, customer-industries face higher prices of their inputs and therefore need to also increase their own

[34] Along this section we assume both $G_i = 0$ and $dG = 0$: no change in government spending and taxes are financed with a negative lump sum transfer $T$ which behaves as a tax deduction.

[35] Thanks to Cobb-Douglas functional form assumption, $(I_n - \rho \cdot A)^{-1} \cdot \boldsymbol{\alpha}^l = \mathbf{1}_n$

price, thus triggering a cascade effect which moves downstream from the top of the production network. This mechanism is also illustrated in our theoretical setting by the expression:

$$d \log \boldsymbol{p} = \frac{\tau}{1-\tau} \cdot H \cdot \mathbf{1}_n \cdot d \log \tau, \tag{11}$$

where $d \log \boldsymbol{p}$ represents the vector of price changes. Notice that, prices change only in response to a tax shock. Basically, in our setting, a tax shock is the analogue of a productivity shock in Acemoglu, Akcigit, and Kerr (2016) and Carvalho (2014): it is a supply side shock which generates spillovers that trickle down to the bottom of the supply chain via production network through the price mechanism.

## C.2 Network effect of a spending shock

Now let's move to government spending shocks and assume that both $\tau = 0$ and $d\tau = 0$. Except for the inclusion of parameter $\rho$ and a slightly different notation, the following derivations are one to one found in Acemoglu, Akcigit, and Kerr (2016); we repeat them for the sake of clarity of the exposition. After log-differentiating the equations that characterize the equilibrium of the model described above, we obtain the following expression:

$$d \log y_i = -\frac{\beta_i}{1+\lambda} \cdot \sum_{j=1}^{n} \cdot \frac{p_j \cdot y_j}{p_i \cdot y_i} \cdot d\tilde{G}_j + \underbrace{\rho \cdot \sum_{j=1}^{n} \hat{a}_{ji} \cdot d \log y_j}_{\text{upstream spatial variable}} + d\tilde{G}_i, \tag{12}$$

where $\tilde{G}_i := \frac{G_i}{y_i}$ and $\hat{a}_{ji} := \frac{x_{ji}}{y_i} = a_{ji} \frac{p_j y_j}{p_i y_i}$.
We can rewrite equation (13) in a compact matrix form:

$$\underset{n\times 1}{d \log \mathbf{y}} = \rho \cdot \underset{n\times n}{\hat{A}^T} \cdot \underset{n\times 1}{d \log \mathbf{y}} + (I_n + \underset{n\times n}{\tilde{\boldsymbol{\Lambda}}}) \cdot \underset{n\times 1}{d\tilde{\mathbf{G}}}, \tag{13}$$

where:

$$\tilde{\mathbf{\Lambda}} = \begin{bmatrix} -\dfrac{\beta_1}{1+\lambda} & -\dfrac{\beta_1}{1+\lambda}\cdot\dfrac{p_2\cdot y_2}{p_1\cdot y_1} & \cdots & -\dfrac{\beta_1}{1+\lambda}\cdot\dfrac{p_n\cdot y_n}{p_1\cdot y_1} \\ -\dfrac{\beta_2}{1+\lambda}\cdot\dfrac{p_1\cdot y_1}{p_2\cdot y_2} & -\dfrac{\beta_2}{1+\lambda} & \cdots & -\dfrac{\beta_2}{1+\lambda}\cdot\dfrac{p_n\cdot y_n}{p_2\cdot y_2} \\ \vdots & \vdots & \ddots & \vdots \\ -\dfrac{\beta_n}{1+\lambda}\cdot\dfrac{p_1\cdot y_1}{p_n\cdot y_n} & -\dfrac{\beta_n}{1+\lambda}\cdot\dfrac{p_2\cdot y_2}{p_n\cdot y_n} & \cdots & -\dfrac{\beta_n}{1+\lambda} \end{bmatrix}$$

and:

$$\hat{A} = \begin{bmatrix} a_{11} & a_{12}\cdot\frac{p_1\cdot y_1}{p_2\cdot y_2} & \cdots & a_{1n}\cdot\frac{p_1\cdot y_1}{p_n\cdot y_n} \\ a_{21}\cdot\frac{p_2\cdot y_2}{p_1\cdot y_1} & a_{22} & \cdots & a_{2n}\cdot\frac{p_2\cdot y_2}{p_n\cdot y_n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1}\cdot\frac{p_n\cdot y_n}{p_1\cdot y_1} & a_{n2}\cdot\frac{p_n\cdot y_n}{p_2\cdot y_2} & \cdots & a_{nn} \end{bmatrix} = A\odot \underbrace{\begin{bmatrix} 1 & \frac{p_1\cdot y_1}{p_2\cdot y_2} & \cdots & \frac{p_1\cdot y_1}{p_n\cdot y_n} \\ \frac{p_2\cdot y_2}{p_1\cdot y_1} & 1 & \cdots & \frac{p_2\cdot y_2}{p_n\cdot y_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{p_n\cdot y_n}{p_1\cdot y_1} & \frac{p_n\cdot y_n}{p_2\cdot y_2} & \cdots & 1 \end{bmatrix}}_{S} = A\odot S$$

where the last identity is meant to underscore the connection between $\hat{A}$ and $A$, the standard input-output matrix, and $S$ represents a scaling matrix.[36]
As done for the case of taxes, we can rewrite equation (13) in its closed form expression:

$$\underset{n\times 1}{d\log\mathbf{y}} = \underbrace{(I_n - \rho\cdot\hat{A}^T)^{-1}}_{\text{upstream propagation}} \cdot (I_n + \tilde{\mathbf{\Lambda}})\cdot \underset{n\times 1}{d\tilde{\mathbf{G}}} \qquad (14)$$

where $\tilde{\mathbf{\Lambda}}$ represents second order GE effects which come from the fact that the government budget constraint holds. The latter term is neglected in our specification since a fiscal consolidation meant to reduce the deficit will not translate into a contemporaneous tax/subsidy (i.e. $T$) reduction.
Equation (14), tells us that the sectoral propagation of a spending shock is driven by the elements in the columns of $\hat{A}$, which describe a sector's sales to other industries. For instance, when $G_i$ decreases, sector $i$ faces a negative demand shock, and reacts by contracting its output and by purchasing less input:

[36] where $*$ denotes the Hadamard product, or element wise product.

those sectors which are suppliers of input to sector $i$, are negatively affected and also shrink their output and purchase less input, and so on and so forth. This type of spillovers represents the aforementioned upstream propagation of demand side shocks.