Uncovering Spatial Productivity Centers using Asymmetric Bidirectional Spillovers

Anthony J. Glass*, Karligash Kenjegalieva† and Mustapha Douch‡,§

December 2019

Abstract

The principal contribution of this paper is to present the first method to sift through a large number of firms in an industry to uncover which firms act as large spatial total factor productivity (TFP) growth centers. We define a large spatial TFP growth center as a firm that is a large net generator of spatial TFP growth spillovers, i.e., it is a source of large TFP growth spill-outs to other firms vis-à-vis the size of the TFP growth spill-ins that permeate to the firm from other firms. We use this definition because, other things being equal, firms would want to locate near a firm that is a net generator of TFP growth spillovers. In the process of presenting the above method we make three further contributions, two of which are methodological and the other relates to our application. First, rather than follow the literature on spatial frontier modeling by considering spatial interaction between firms in a single network, we introduce a more sophisticated model that is able to account for spatial interaction in multiple networks. Second, we obtain bidirectional spatial TFP growth decompositions by complementing a unidirectional decomposition in the literature, where the spillover components are spill-ins to a firm, with a decomposition that includes spill-out components. Third, from a spatial revenue frontier for U.S. banks ($1998-2015$), we find a number of cases where banks that represent large spatial TFP growth centers have branches that cluster together, while in several states we find no such clusters.

Key words: Productivity and competitiveness; Spatial stochastic frontier analysis; Revenue function; Multiple spatial networks; U.S. banks.

*School of Business and Economics and Centre for Productivity and Performance, Loughborough University, Loughborough, Leics, UK, LE11 3TU. Email: a.j.glass@lboro.ac.uk

†School of Business and Economics and Centre for Productivity and Performance, Loughborough University, Loughborough, Leics, UK, LE11 3TU. Email: k.a.kenjegalieva@lboro.ac.uk

‡Aston University Business School, Birmingham, B4 7ET, UK. Email: m.douch@aston.ac.uk

§Corresponding author

1 Introduction

There has been a lot of methodological and applied non-spatial studies on the decomposition of total factor productivity (TFP) growth (Diewert and Fox, 2017, Sun *et al.*, 2015, Orea, 2002, Oude Lansink *et al.*, 2015, and Tsionas and Mallick, 2019, to name but a small selection of studies). There is, of course, variation in the precise forms of the decompositions in this literature, but a typical decomposition of a firm's TFP growth may consist of returns to scale change, technical/economic efficiency change and technological change. Building on the more general parametric literature that models productivity spillovers (e.g., Chandra and Staiger, 2007, Takii, 2005, and Girma and Wakelin, 2007), an emerging parametric literature sets out the methodology to augment the decomposition of a firm's own TFP growth with the decomposition of the TFP growth spillovers it receives from other firms (Glass *et al.*, 2013, and Glass and Kenjegalieva, 2019, where the latter is denoted as G&K from hereon). As this literature exclusively focuses on the TFP growth spillovers that gravitate to a firm, our first methodological contribution is to propose a new spatial decomposition of TFP growth by augmenting the decomposition of a firm's own TFP growth with the decomposition of the TFP growth spillovers it transmits to other firms. By conjoining this contribution with the aforementioned emerging literature we obtain asymmetric bidirectional spatial decompositions of TFP growth.

Using a spatial stochastic frontier model, which in terms of spatial variables contains at least a spatial lag of the dependent variable, otherwise known as a spatial autoregressive (SAR) variable, we obtain our new spatial decomposition of TFP growth. In the spatial stochastic frontier literature the models only consider spatial interaction within a single network (Druska and Horrace, 2004; G&K; Glass *et al.*, 2013; 2016; 2018a; 2019; Tsionas and Michaelides, 2016; Gude *et al.*, 2018; Orea *et al.*, 2018; Horrace *et al.*, 2019). Our second methodological contribution extends this literature by introducing a spatial stochastic frontier model that caters for multiple spatial networks via multiple spatial lags of the dependent variable, which we then use to obtain our new spatial decomposition of TFP growth. For firms this will often be a better representation of what is actually the case because, in practice, a firm may simultaneously belong to a number of different networks.

Our third and principal contribution is to conjoin the asymmetric bidirectional measures of TFP growth spillovers to introduce a methodology to uncover firms that act as spatial TFP growth centers in an industry. We define a firm as being a spatial TFP growth center if the sum of the TFP growth spillovers it transmits to other firms is greater than the sum of the spillovers it receives. We use this definition of a spatial TFP growth center because, everything else being equal, a firm would want to benefit from the net generation of TFP growth spillovers by locating near a center. As we also decompose the asymmetric bidirectional measures of TFP growth spillovers into asymmetric bidirectional spillovers of technological change, returns to scale change and efficiency change, these decompositions indicate the components that are key in determining which firms act as spatial TFP growth centers in an industry. As firm level data samples are typically large there are likely to be a large number of firms in empirical applications who are net generators and net recipients of TFP growth spillovers. The issue then is which firms are the largest net generators in a sample and which are largest net recipients.

Our fourth and final contribution is to provide an empirical application of our methodological advances to the U.S. banking industry. We complement G&K's spatial cost frontier based application of their spatial TFP growth decomposition by focusing on the income side of the industry. This involves estimating a spatial stochastic revenue frontier with multiple networks for the period $1998 - 2015$. By using a translog functional form we calculate elasticities outside the sample mean, which we use to obtain our new annual spatial TFP growth decomposition. This annual decomposition plays a key role in uncovering which U.S. banks act as spatial TFP growth centers in the industry. In particular, we use our annual decomposition to: (i) uncover which banks represent spatial growth centers in 2015, as the final year of our analysis is the most relevant for future policy making; and (ii) investigate if there are differences between these results for 2015 and those for two subperiods- pre-crisis ($1998-2007$) and crisis and beyond ($2008-2015$).

It is informative for bank regulatory authorities to know which banks represent the largest spatial TFP growth centers and which banks are the largest net recipients of TFP growth spillovers. This is because this information provides new insights into which banks play key roles in the productivity effects that emanate from the linkages between banks. Such insights for policy makers suggests there is scope for wider application of our methodology to other industries. Although our application is to banks across the U.S., our methodology to uncover spatial TFP growth centers can also be applied to firms that cluster in a particular location (e.g., the hi-tech enterprises in the Santa Clara Valley area of California and the large scale manufacturing firms in the Pearl Delta area of China).

Our quantitative analysis of productivity centers is the first of its type. The two closest related literatures to our analysis are the quantitative studies of hubs and a study by Pesaran and Yang (2019) that analyzes how the sectors that make up a country's production network affect its aggregate output. We now turn to briefly discuss these two literatures and at the outset point out that in network theory a hub-and-spoke network has a particular network structure, whereas here we are not so prescriptive about the structure of bank branch networks. Irrespective of the particular form of a bank branch network, we simply use banks that have overlapping branch networks to uncover the banks that act as productivity centers.

There are two broad strands of the quantitative literature on hubs. The first strand is a large methodological OR literature that analyzes the optimal locations of hubs in transportation and logistics (T&L) networks (e.g., Alumur and Kara, 2008, Campbell and O'Kelly, 2012, Contreas, *et al.*, 2011, and Ishfaq and Sox, 2011). The second adopts a different approach because, rather than propose the location of a hub, it analyzes the effect of hubs on a particular variable when the hubs can be selected prior to the empirical modeling as they are well-known. Examples of the latter analyze the effect of hub-and-spoke airline networks on fares (Brueckner *et al.*, 1992; Borenstein, 1989) and airport congestion (Mayer and Sinai, 2003), and the effect of universities as research hubs on the wider technological innovation in industries and surrounding areas (Anselin *et al.*, 1997, Cohen *et al.*, 2002, and Fromhold-Eisebith and Werker, 2013, to name but a few studies from this large literature). Our paper does not fit into either strand of this quantitative literature on hubs because we do not consider T&L networks, or analyze the effect of productivity centers that are easily observed and can therefore be selected in advance of the modeling, which is not a common situation. Instead, the new line of enquiry we pursue considers the case of a large number of firms in an industry when the firms that act as spatial economic performance centers are not easily observed, which is often the case in practice. For this type of case we propose a resolution by setting out the first method to sift through a large number of firms in an industry to uncover which firms act as spatial TFP growth centers (i.e., those firms that are large net generators of TFP growth spillovers). We use spatial TFP growth as our measure of spatial economic performance because TFP is widely recognized as a comprehensive measure that incorporates both demand and supply phenomena.

There are parallels between our analysis and Pesaran and Yang (2019), but whereas our approach operates at the micro level of a large number of firms, their method focuses on the macro level. In their model sector-specific shocks affect the aggregate output of a country if there are "dominant" sectors, where a sector is dominant if its outdegree is not bounded by the number of sectors in the country. Using information from an input-output table, they define the outdegree of a sector as the share of the sector's output that is used as intermediate inputs by all the other sectors in the country. They then set out how to use this information to estimate an exponent that measures the degree of dominance (or pervasiveness) of a sector in the determination of a country's aggregate output. In particular, this exponent controls the rate at which the outdegree of a sector rises with the number of sectors in the country.

In the context of our micro setting, the above outdegree definition may not be appropriate for many industries because firms may not use other firms' outputs as intermediate inputs, and in industries where they do, data on this interfirm economic interaction may not be in the public domain. For example, banks interact in this way through interbank lending, but data on such lending is not publicly available. This necessitates therefore that we use our alternative approach.

By way of an insight into our key empirical findings, based on our analysis of 2015, $1998-2007$ and $2008-2015$, we find that the banks which act as large spatial TFP growth centers tend to have branches that cluster together in the same areas throughout our study period. In line with our expectations, these areas are the northeast and cities elsewhere that are among the largest economies (e.g., Los Angeles and Chicago). Over our study period we observe an increase in the clustering of the branches of the banks that act as large spatial TFP growth centers, which indicates that these clusters are becoming an increasingly prominent feature of the industry. That said, we find no evidence of banks that act as large spatial TFP growth centers having clusters of branches in a number of contiguous states (Idaho; Montana; North Dakota; South Dakota; Wyoming; Maine; Vermont; and New Hampshire).

Additionally, it is important to note that the basic principle of our method to uncover spatial TFP growth centers is sufficiently general that it can be applied to any spatial measure. For example, we could apply our method to uncover firms that act as spatial technical innovation centers in a sample of hi-tech firms. A firm would be classified as a spatial technical innovation center if the sum of the technical innovation spillovers it transmits to other firms are greater than the sum of the spillovers it receives. The results from such an analysis could be used to improve the efficacy of public innovation policies by targeting policies at firms that represent spatial technical innovation centers.

Specifically, we estimate a multi-network spatial Durbin stochastic frontier, where the spatial interaction between the banks in the multiple networks is accounted for by multiple sets of spatial lags of the dependent and independent variables. In addition to, as we have noted above, the multi-network aspect of the model being more representative of the links between U.S. banks, we estimate this model specification for three further reasons, which we now turn our attention to.

First, although a spatial error/inefficiency specification (which contains a spatial lag of the disturbance/inefficiency) and SAR and spatial Durbin models all account for global spatial dependence (i.e., 1st order and higher order neighbor spatial interaction), as we discuss in more detail in due course, with the spatial error/inefficiency specification the only spillover elasticity in the model relates to the disturbance/inefficiency.[1] In other words, with the spatial error/inefficiency specification, the only type of spillover that is modeled is the impact on a firm's dependent variable from the spillover of other firms' disturbances/inefficiencies. From the SAR and spatial Durbin specifications, however, as well as being able to obtain for our new spatial TFP growth decomposition the efficiency spillovers to (from) a firm from (to) other firms, in contrast to a spatial error/inefficiency model and in line with the requirements of our new decomposition, the spillover elasticities can also be related to the independent variables. As a result of this property of the SAR and spatial Durbin specifications, we obtain the two further spillover components for our decomposition (growth in technological change spillovers and growth in returns to scale spillovers). Second, we favor the spatial Durbin specification over the SAR model because the

[1] For the further discussion of the spillover elasticity from a spatial error/inefficiency specification see footnote 7 in subsection 2.3.

former includes spatial lags of the independent variables, while the latter omits these lags. The inclusion of these spatial lags in the spatial Durbin specification accounts for local spatial dependence. This is a further type of spatial dependence that is often found to be present in empirical applications and relates to only 1st order spatial interaction between the independent variables of neighboring firms. Third, from the SAR model the ratio of the own and spillover elasticities (referred to in the spatial literature as direct and indirect elasticities) is the same for all the independent variables, which is implausible. This is not the case though with the spatial Durbin specification because of the presence of the spatial lags of the independent variables in the model.

The remainder of this paper is organized as follows. Section 2 has three parts. In the first part we provide a general presentation of the structural form of our new spatial Durbin stochastic frontier model. The second part sets out our approach to estimate the parameters of the structural model, the own (i.e., non-spatial) time-invariant and time-varying efficiencies, and the combined measure that draws together these efficiencies. In the third part, as the structural form of our model does not yield interpretable global spillovers, we present how we calculate a set of elasticities from the reduced form model (direct, indirect spill-in and spill-out, and total), and then layout our method to compute the corresponding efficiencies. In section 3 we set out our methodology to uncover the firms that act as spatial TFP growth centers in an industry, which first involves presenting our method to obtain the asymmetric spatial TFP growth decompositions. In section 4 we present the application to U.S. banks and in section 5 we conclude.

2 A Spatial Stochastic Frontier Model with Multiple Spatial Networks

2.1 General Presentation of the Structural Form of the Model

Before we formally present and discuss our model in detail, we provide an intuitive overview of its new methodological feature. This feature extends the literature on spatial stochastic frontier models that consider interactions between firms within a single spatial network (via a single spatial lag of the dependent variable), to the more complex setting of multiple simultaneous spatial networks (via multiple spatial lags of the dependent variable). The motivation for this extension is to represent the different types of spatial linkages that simultaneously exist between firms that are in multiple networks, where these networks may represent the different business segments of firms. In our model there are M spatial networks that are indexed $m \in 1, ..., M$. When we apply our model we favor a reasonably small M that relates to reasonably distinct spatial networks. This does not detract from our model as it guards against collinearity between some of spatial lags of the dependent variable when M is quite large. With this in mind, in our application we use just two spatial weights matrices, which reflect the linkages between the banks across two networks of very different branch types. To obtain the small M we construct one spatial weights matrix by aggregating a number of branch types across a broader category. In further applications of our model it may be necessary to aggregate networks using a similar approach.

To emphasize the generality of our model and estimator, the following presentation is for the general case of a concave frontier rather than a specific concave frontier technology (e.g., a production, profit or revenue frontier). Our model and estimator, however, is in no way constrained to a concave frontier and can easily be adapted to convex frontier technologies (e.g., a cost frontier). The general structural form of our panel data spatial Durbin stochastic frontier (SDSF) model with multiple simultaneous spatial networks is given in Eq. 1, where the variables are logged.

$$y_{it} = \alpha + x'_{it}\beta + \begin{pmatrix} \sum_{j=1}^{N} w_{ij1}x'_{jt}\vartheta_1 + ... \\ + \sum_{j=1}^{N} w_{ijM}x'_{jt}\vartheta_M \end{pmatrix} + \begin{pmatrix} \delta_1 \sum_{j=1}^{N} w_{ij1}y_{jt} + ... \\ + \delta_M \sum_{j=1}^{N} w_{ijM}y_{jt} \end{pmatrix} + \varrho_i + v_{it} - \eta_i - u_{it}. \tag{1}$$

The panel data comprises observations for N firms over T time periods, which are indexed $i, j \in 1, ..., N$ $\forall\ i, j$ and $t \in 1, ..., T$. y_{it} is the observation for the dependent variable (output, profit or revenue) for the ith firm in period t; α is the common intercept; x'_{it} is the $(1 \times K)$ vector of observations for the non-spatial regressors which are indexed $k \in 1, ..., K$; and β is the associated vector of regression parameters. x_{it} includes the variables which together with y_{it} represent the frontier technology.

The M spatial networks in our model represent different sources of spillovers and correspond to the M $(N \times N)$ spatial weights matrices, which are indexed $\mathbf{W_m} \in \mathbf{W_1}, ..., \mathbf{W_M}$. One must specify $\mathbf{W_1}, ..., \mathbf{W_M}$ in advance of estimating the model, where the elements of $\mathbf{W_m}$ are the non-negative constant $i, j-$th spatial weights w_{ijm}. For the mth spatial network $\mathbf{W_m}$ represents the spatial arrangement of the firms and the strength of the spatial linkages between the firms. All the elements on the main diagonal of $\mathbf{W_m}$ are therefore set to zero because a firm cannot be spatially linked to itself in a network. As $\mathbf{W_m}$ is exogenous it is often specified using some measure of geographical proximity. Although this is also the case in our empirical application to U.S. banks we adopt an innovative approach. This involves constructing for each branch type a spatial weights matrix that represents the geographical overlap of bank branch networks.

Once $\mathbf{W_1}, ..., \mathbf{W_M}$ have been specified we can construct the two groups of spatial variables in our model. The first group is the M spatial lags of the dependent variable, $\sum_{j=1}^{N} w_{ij1}y_{jt}, ..., \sum_{j=1}^{N} w_{ijM}y_{jt}$. These spatial lags capture global spatial dependence in the M networks and are endogenous, which our estimator accounts for. The associated parameters $\delta_1, ..., \delta_M \in \left[\frac{1}{\min(f_1^{\min}, ..., f_M^{\min})}, \frac{1}{\max(f_1^{\max}, ..., f_M^{\max})}\right]$, where this space is sufficiently general and $f_m^{\max}$ is the most positive real characteristic root of $\mathbf{W_m}$. Specifically, $\mathbf{W_m}$ denotes a normalized specification of the mth spatial weights matrix, where the normalization we use in our empirical application gives $f_1^{\max}, ..., f_M^{\max} = 1$. For details of our approach to this normalization see subsection 4.1. In our application and as is common in the empirical spatial literature $\mathbf{W_m}$ before normalization is asymmetric and may therefore have complex roots after normalization. For the case of complex roots and using our notation, LeSage and Pace (2009) (LSP from hereon) prove that the lower limit of δ_m is the inverse of $f_m^{\min}$, where $f_m^{\min}$ is the most negative real characteristic root of $\mathbf{W_m}$.

The second group of spatial variables in our model are the M exogenous spatial lags of x'_{it} (i.e., $\sum_{j=1}^{N} w_{ij1}x'_{jt}, ..., \sum_{j=1}^{N} w_{ijM}x'_{jt}$), where $\vartheta_1, ..., \vartheta_M$ are the associated vectors of parameters. These spatial lags capture local spatial dependence in the M networks and together with the spatial lags of the dependent variable shift the frontier technology. If we were to omit the M spatial lags of x'_{it} then Eq. 1 becomes the SAR model. We do not omit these variables because often local spatial variables are found to be important determinants in the empirical spatial literature. Due to the inclusion of these local spatial variables our model belongs to the class of spatial Durbin models.

The error structure in our model is $\widetilde{\varepsilon}_{it} = \varepsilon_i + \varepsilon_{it} = \varrho_i + v_{it} - \eta_i - u_{it}$, where $\varepsilon_i = \varrho_i - \eta_i$ is the time-invariant error and $\varepsilon_{it} = v_{it} - u_{it}$ is the time-varying error. To distinguish between the four error components our estimator relies on the components being independently distributed. We do not therefore use firm fixed effects to account for unobserved heterogeneity because although they will be correlated with the regressors which is appealing, the time-varying idiosyncratic errors will not be independently distributed as they will also be correlated with the fixed effects.

G&K model unobserved firm heterogeneity using random effects which are independently distributed because, as is standard, these effects are random errors. As a result, random effects are uncorrelated with the regressors, which may not be the case in many applications. We therefore follow the spatial stochastic frontier model with a single spatial network in Glass *et al.* (2018a) and also Debarsy (2012) and Wang and Lee (2013) in the spatial literature where there is no inefficiency component(s), by modeling unobserved firm heterogeneity using common correlated random effects. This approach allows the effects to be correlated with the regressors while maintaining independently distributed error components and involves specifying a firm common correlated random effect as

$$\varrho_i = \overline{x}_i'\zeta + \sum_{j=1}^{N} w_{ij1}\overline{x}_j'\xi_1 + ... + \sum_{j=1}^{N} w_{ijM}\overline{x}_j'\xi_M + \kappa_i. \quad (2)$$

$\kappa_i \sim N(0, \sigma_\kappa^2)$ is a random error and therefore represents the component of the firm specific effect that is uncorrelated with the regressors. $\overline{x}_i' = \frac{1}{T}\sum_{t=1}^{T} x_{it}'$ and $\sum_{j=1}^{N} w_{ijm}\overline{x}_j' = \frac{1}{T}\sum_{t=1}^{T}\sum_{j=1}^{N} w_{ijm}x_{jt}'$, where these means account for correlation between the firm specific effects and each of the regressors. Associated with these means are the vectors of parameters ζ and $\xi_1, ..., \xi_M$.

Turning to the inefficiency components of the error structure, $\eta_i \sim N^+\left(0, \sigma_\eta^2\right)$ is own time-invariant inefficiency and $u_{it} \sim N^+\left(0, \sigma_u^2\right)$ is own time-varying inefficiency. Here η_i and u_{it} are assumed to have half-normal distributions, which is common in the stochastic frontier literature, although our estimation procedure is sufficiently general to accommodate alternative distributional assumptions. The own inefficiencies from Eq. 1 are directly comparable to the inefficiencies from the corresponding non-spatial frontier model, and following the spatial literature (e.g., Anselin, 2003) Eq. 1 represents the structural form of our model as it includes the M SAR variables. In the reduced form model, which is used to compute the elasticities and efficiency spillovers, the SAR variables do not feature.

One can test the appropriateness of the error structure in Eq. 1 for an empirical application using the one-sided likelihood ratio test in Gouriéroux *et al.* (1982). This involves performing the test for each of the error components, where the asymptotic distribution of the test statistic is a mixture of chi-squared distributions, $\frac{1}{2}\chi_0^2 + \frac{1}{2}\chi_1^2$. For $G \in \{\varrho, v, \eta, u\}$ rejection of the null $\left(\widehat{\sigma}_G^2 = 0\right)$ in favor of the alternative hypothesis $\left(\widehat{\sigma}_G^2 > 0\right)$ constitutes evidence of the presence of the error component. Since both of the nulls for the inefficiency components may not be rejected in an empirical application, our model has the advantage of nesting the corresponding SDSF models that consider only time-invariant or time-varying inefficiency. G&K use the same test for the firm specific random effect in their error structure, which is akin to testing for the presence of the κ_i component in ϱ_i in our model. With our model, in addition to applying this test to ϱ_i, we can go a step further by applying the test to the three types of component that make up ϱ_i, i.e., κ_i, $\overline{x}_i'$ and $\sum_{j=1}^{N} w_{ijm}\overline{x}_j'$. A significant κ_i component indicates that part of ϱ_i is uncorrelated with the regressors. Significant $\overline{x}_i'$ and/or $\sum_{j=1}^{N} w_{ijm}\overline{x}_j'$ points to x_{it}' and/or $\sum_{j=1}^{N} w_{ijm}x_{jt}'$ being significantly correlated with ϱ_i.

2.2 Model Estimation and the Own Efficiency Measures

We estimate the model in Eq. 3 as we treat $\overline{x}_i'$ and $\sum_{j=1}^{N} w_{ij1}\overline{x}_j', ..., \sum_{j=1}^{N} w_{ijM}\overline{x}_j'$ as auxiliary regressors.

$$
y_{it} = \alpha + x'_{it}\beta + \begin{pmatrix} \sum_{j=1}^{N} w_{ij1}x'_{jt}\vartheta_1 + ... \\ + \sum_{j=1}^{N} w_{ijM}x'_{jt}\vartheta_M \end{pmatrix} + \begin{pmatrix} \delta_1 \sum_{j=1}^{N} w_{ij1}y_{jt} + ... \\ +\delta_M \sum_{j=1}^{N} w_{ijM}y_{jt} \end{pmatrix} + \overline{x}'_i\zeta +
$$

$$
\sum_{j=1}^{N} w_{ij1}\overline{x}'_j\xi_1 + ... + \sum_{j=1}^{N} w_{ijM}\overline{x}'_j\xi_M + \kappa_i + v_{it} - \eta_i - u_{it}, \qquad (3)
$$

where now the time-invariant and time-varying errors are $\varepsilon_i = \kappa_i - \eta_i$ and $\varepsilon_{it} = v_{it} - u_{it}$.

Given Eq. 3 is more complex than a simpler spatial stochastic frontier model with a single SAR variable and only one type of inefficiency (time-varying or time-invariant), we use a practical multi-step maximum likelihood (ML) estimation procedure.[2] This procedure comprises three steps, where in step 1 we estimate the SAR parameters $\delta_1, ..., \delta_M$, the intercept α and the parameters on the other variables (i.e., β, $\vartheta_1, ..., \vartheta_M$, ζ and $\xi_1, ..., \xi_M$). In steps 2 and 3 we estimate u_{it} and η_i, respectively. We now present the salient features of our estimation procedure and in the Appendix we provide the technical details of the log-likelihood function(s) for each step. To proceed we invoke the standard type of reparameterization of Eq. 3 that is adopted in the stochastic frontier literature (e.g., Aigner *et al.*, 1977). This involves using $\sigma^2_{\eta\kappa} = \sigma^2_\eta + \sigma^2_\kappa$ and $\lambda_{\eta\kappa} = \sigma_\eta/\sigma_\kappa$, and $\sigma^2_{uv} = \sigma^2_u + \sigma^2_v$ and $\lambda_{uv} = \sigma_u/\sigma_v$. It follows that $\sigma^2_\eta = \sigma^2_{\eta\kappa}/\left(1 + \lambda^2_{\eta\kappa}\right)$; $\sigma^2_\kappa = \sigma^2_{\eta\kappa}\lambda^2_{\eta\kappa}/\left(1 + \lambda^2_{\eta\kappa}\right)$; $\sigma^2_u = \sigma^2_{uv}/\left(1 + \lambda^2_{uv}\right)$; and $\sigma^2_v = \sigma^2_{uv}\lambda^2_{uv}/\left(1 + \lambda^2_{uv}\right)$.

We estimate the aforementioned parameters in step 1 by transforming the spatial stochastic frontier model in Eq. 3 into the spatial model in Eq. 4. As a result of this transformation firms are not yet permitted to be inefficient. In subsequent steps this is relaxed to estimate the time-varying and time-invariant inefficiencies. To rewrite Eq. 3 as Eq. 4 we transform the intercept and the negatively skewed time-invariant and time-varying errors ε_i and ε_{it} as follows: $\alpha^\circ = \alpha - \mu_{\varepsilon_i} - \mu_{\varepsilon_{it}}$; $\varepsilon^\circ_i = \varepsilon_i + \mu_{\varepsilon_i} = \kappa_i - \eta_i + \mu_{\varepsilon_i}$; and $\varepsilon^\circ_{it} = \varepsilon_{it} + \mu_{\varepsilon_{it}} = v_{it} - u_{it} + \mu_{\varepsilon_{it}}$, where $\mu_{\varepsilon_i} = \mathrm{E}(\eta_i)$ and $\mu_{\varepsilon_{it}} = \mathrm{E}(u_{it})$. As a result, the transformed time-invariant and time-varying error components satisfy the zero-mean condition, i.e., $\varepsilon^\circ_i \sim N(0, \sigma^2_{\varepsilon^\circ_i})$ and $\varepsilon^\circ_{it} \sim N(0, \sigma^2_{\varepsilon^\circ_{it}})$.[3]

[2] This practicality is the result of breaking down the estimation problem into steps, which faciltates in each step convergence of the log-likelihood function(s). Although this is beneficial, the cost associated with this practicality is the loss of some statistical efficiency of the parameter and efficiency estimates from stepwise estimation of the model (vis-à-vis estimation in a single step). See Gude *et al.* (2018) and Horrace *et al.* (2019) for ML methods that estimate spatial stochastic frontier models all in a single step. Their models, however, consider one spatial network and only time-varying inefficiency and are thus much simpler than our model, which is why we use a practical stepwise estimator.

[3] Our estimator extends to multiple spatial networks the multi-step ML estimator in Kumbhakar *et al.* (2014) for the corresponding random effects non-spatial stochastic frontier model (i.e., our model with the means of the x variables and their spatial lags and the spatial lags of x and y omitted). Simulated ML has also been used by Filippini and Greene (2016) to estimate in a single step the corresponding non-spatial stochastic frontier model with common correlated random effects (i.e., our model with the spatial lags of the means of the x variables and the spatial lags x and y omitted). Since the focus here is on presenting our method to uncover firms that act as spatial TFP growth centers and given ML methods are well-established approaches to estimate corresponding non-spatial stochastic frontier models, we do not formally pursue the asymptotic properties of our estimator. It is clear though that the asymptotic properties of our estimator in the first step (e.g., the consistency of the estimates of the parameters in Eq. 4) would not raise any issues. This is because these properties would be the amalgamation of those for the pooled least squares estimator of the non-spatial model with common correlated random effects (Pesaran, 2006), and the estimator of the SAR random effects model using ML (Lee and Yu, 2012). It should also be noted that in the same way as the x variables in Eq. 4 must be uncorrelated with ε°_i and ε°_{it} to obtain consistent estimates of the coefficients on these variables, the means of the x variables and their spatial lags must also be uncorrelated with these error components to obtain consistent estimates of the coefficients on these auxiliary regressors.

$$y_{it} = \alpha^{\circ} + x'_{it}\beta + \begin{pmatrix} \sum_{j=1}^{N} w_{ij1}x'_{jt}\vartheta_1 + ... \\ + \sum_{j=1}^{N} w_{ijM}x'_{jt}\vartheta_M \end{pmatrix} + \begin{pmatrix} \delta_1 \sum_{j=1}^{N} w_{ij1}y_{jt} + ... \\ +\delta_M \sum_{j=1}^{N} w_{ijM}y_{jt} \end{pmatrix} + \overline{x}'_i\zeta +$$
$$\sum_{j=1}^{N} w_{ij1}\overline{x}'_j\xi_1 + ... + \sum_{j=1}^{N} w_{ijM}\overline{x}'_j\xi_M + \varepsilon_i^{\circ} + \varepsilon_{it}^{\circ}. \tag{4}$$

To give spatial econometrics greater economic foundation Corrado and Fingleton (2012) suggest that more emphasis should be placed on using spatial techniques to estimate equations that are well grounded in economic theory. We can also apply this research direction to address a particular issue in spatial econometrics because such an approach can alleviate a potential problem with the identification of the spatial parameters in Eq. 4, i.e., the coefficients on both the local spatial variables $\left(\sum_{j=1}^{N} w_{ij1}x'_{jt}, ..., \sum_{j=1}^{N} w_{ijM}x'_{jt}\right)$ and global spatial variables $\left(\sum_{j=1}^{N} w_{ij1}y_{jt}, ..., \sum_{j=1}^{N} w_{ijM}y_{jt}\right)$. Identification of these spatial parameters can be problematic if x'_{it} suffers from omitted variables, as there is a tendency for the effect of these omitted variables to inflate the spatial parameters because the omitted variables are spatially correlated. To circumvent this problem in our empirical application y_{it} and x'_{it} are specified according to a well-established function from production theory, i.e., a translog revenue function, where revenue is a function of first and second order output prices and input quantities and interactions between these first order variables. Consequently, in our empirical model the local and global spatial parameters are identified because according to economic theory no variables are omitted from x'_{it}.

Using Q to denote the Hessian matrix associated with the log-likelihood function for the model in Eq. 4, it is well known that the variance-covariance matrix to calculate the standard errors of the parameters is $-Q^{-1}$. Calculating this Hessian, however, is not as straightforward as for the corresponding non-spatial model. To calculate the Hessian we follow the approach of LSP for spatial non-frontier models, which involves using a mixture of analytical and numerical methods. All the second order derivatives that form the Hessian are computed analytically with the exception of $\partial^2 LL/\partial\delta_1^2, ..., \partial^2 LL/\partial\delta_M^2$, which are calculated numerically. Calculating the second order derivatives of a log-likelihood function analytically rather than numerically is far less sensitive to badly scaled data. Calculating $\partial^2 LL/\partial\delta_1^2, ..., \partial^2 LL/\partial\delta_M^2$ numerically when N is large though, as is the case in our empirical application, simplifies matters somewhat as it avoids a number of large matrix multiplications involving $\mathbf{W_1}, ..., \mathbf{W_M}$.

As will become apparent in due course, the spatial multiplier matrix $(\mathbf{I_N} - \delta_1\mathbf{W_1} - ... - \delta_M\mathbf{W_M})^{-1}$ plays a key role in the reduced form of our model as it is where the global spillovers come from, where $\mathbf{I_N}$ is the $(N \times N)$ identity matrix. In our empirical application this matrix inversion does not pose any problems and, as a result, we obtain the spatial multiplier matrix we need to calculate the direct, indirect and total elasticities and efficiencies. See subsection 2.3 for discussion of the role of the spatial multiplier matrix in the calculation of the direct, indirect and total elasticities and efficiencies.

Using $\widehat{\varepsilon}_{it}^{\circ}$ from step 1, in step 2 we begin the process of splitting $\widehat{\varepsilon}_{it}$ into $\widehat{u}_{it}$ and $\widehat{v}_{it}$ by using ML to obtain the estimate of λ_{uv}, which is then used to obtain the estimate of σ_{uv}. See the Appendix for technical details on this. We then also in step 2 use $\widehat{\lambda}_{uv}$, $\widehat{\sigma}_{uv}$ and the relevant other estimates to calculate the consistent estimate of own time-varying inefficiency u_{it} (conditional on ε_{it}), using the following well-established Jondrow *et al.* (1982) (JMLS) estimator:

$$\widehat{u}_{it} = \mathrm{E}\left(u_{it}|\varepsilon_{it}\right) = \frac{\sigma_u\sigma_v}{\sigma_{uv}}\left(\frac{\phi_{it}}{1-\Phi_{it}} - \frac{\varepsilon_{it}\lambda_{uv}}{\sigma_{uv}}\right). \tag{5}$$

$\Phi_{it} = \Phi(\varepsilon_{it}\lambda_{uv}/\sigma_{uv})$, where Φ is the standard normal cumulative distribution function; $\phi_{it} = \phi(\varepsilon_{it}\lambda_{uv}/\sigma_{uv})$, where ϕ is the probability density function for the standard normal distribution; and $\varepsilon_{it} = \varepsilon_{it}^{\circ} - \mu_{\varepsilon_{it}}$.

In step 3 we use the same approach to estimate own time-invariant inefficiency η_i (conditional on ε_i) as we use to calculate $\widehat{u}_{it}$ in step 2. In brief, this involves using $\widehat{\varepsilon}_i^{\circ}$ from step 1 to split $\widehat{\varepsilon}_i$ into $\widehat{\eta}_i$ and $\widehat{\kappa}_i$, by calculating $\widehat{\lambda}_{\eta\kappa}$ and subsequently $\widehat{\sigma}_{\eta\kappa}$ and then $\widehat{\eta}_i$ via the JMLS method. Finally, in a similar way to how we obtain $\widehat{\varepsilon}_{it}$ and $\widehat{\varepsilon}_i$, we scale $\widehat{\alpha}^{\circ}$ to obtain the consistent estimate of the constant term, $\widehat{\alpha} = \widehat{\alpha}^{\circ} + \widehat{\mu}_{\varepsilon_i} + \widehat{\mu}_{\varepsilon_{it}}$.[4]

We can transform the log form in Eq. 1 into its multiplicative form for level data by taking the exponents of the terms in Eq. 1. For the inefficiency terms this is the widely used Battese and Coelli (1988) method to transform an inefficiency estimate into the efficiency measure. For our model this gives estimates of what we refer to as own net time-varying efficiency, $NVE_{it} = \exp(-\widehat{u}_{it})$, and own net time-invariant efficiency, $NIE_i = \exp(-\widehat{\eta}_i)$. We use net to indicate that these efficiencies are net of time-invariance and time-variance, respectively. Based also on the multiplicative form of Eq. 1, we compute the estimate of own combined efficiency, $GVE_{it} = \exp[-(\widehat{\eta}_i + \widehat{u}_{it})] = NIE_i \times NVE_{it}$, which we refer to as own gross time-varying efficiency.[5] In line with simpler multiplicative models that contain only an own time-varying/invariant measure of efficiency, NVE_{it}, NIE_i, $GVE_{it} \in [0, 1]$.

2.3 Bidirectional Elasticities and Efficiency Spillovers

For the corresponding non-spatial form of Eq. 1 and the local spatial form one can interpret the fitted model as the estimated coefficients on the variables are elasticities.[6] As LSP note, however, the fitted structural form of a spatial model, which in terms of spatial variables contains at least one SAR variable, does not yield interpretable spillovers. This is because the spillover elasticities for the x variables from such a model are a function of the SAR parameter(s). Following LSP, who developed the method to interpret such models, we use the fitted parameters for the structural form of our model to calculate direct, indirect and total elasticities, which are partially/entirely made up of a spillover. A direct elasticity is interpreted in the same way as an elasticity from a non-spatial model, although a direct elasticity contains a particular type of spillover known as feedback, which in empirical applications is typically found to be small (e.g., Autant-Bernard and LeSage, 2011). Feedback is the portion of the effect of a change in an independent variable for a particular firm that reverberates back to the same firm's dependent variable, via the impact on the dependent variables of the other firms in the sample.

Indirect elasticities are entirely due to spillovers and are calculated in two ways as they are bidirectional. This gives rise to two interpretations of an indirect elasticity: (i) the effect on a firm's dependent variable when there is a change in an independent variable for all the other firms in the sample; and (ii) the effect on the dependent variables of all the other firms in the sample when there is a change in an independent variable for a particular firm. The mean indirect spill-in and spill-out elasticities across the

[4]In contrast to this upward scaling of $\widehat{\alpha}^{\circ}$ for stochastic production, revenue and profit frontiers, for a stochastic cost frontier $\widehat{\alpha} = \widehat{\alpha}^{\circ} - \widehat{\mu}_{\varepsilon_i} - \widehat{\mu}_{\varepsilon_{it}}$.

[5]In the non-spatial parametric efficiency literature own GVE is referred to as own overall time-varying efficiency. We, however, depart from this diction for non-spatial frontiers as we proceed beyond own efficiency measures. Having computed own GVE and borrowing some terminology from the spatial literature, we compute direct, indirect and total GVE, which are all spatial efficiencies as they are entirely/partially made up of a spillover efficiency. As a result, we avoid referring to total GVE as total overall time-varying efficiency, which would be confusing. As we refer to a combined efficiency as a gross measure, it is logical to refer to the time-invariant and time-varying components of a GVE measure as net efficiencies. Additionally, own NIE, NVE and GVE should not be confused with the net and gross efficiencies in Coelli *et al.* (1999), as the interpretations of net and gross in their model are entirely different to how we interpret net and gross here.

[6]The non-spatial form is Eq. 1 with the M spatial lags of x'_{it}, their means in ϱ_i and the M spatial lags of y_{it} omitted. The local spatial form is Eq. 1 with the M spatial lags of y_{it} omitted.

sample of firms (i.e., (i) and (ii) above, respectively) will be equal, but for an individual firm they will differ in magnitude. Summing the direct and indirect elasticities for a variable gives a total elasticity. For an individual firm the asymmetric bidirectional indirect elasticities lead to asymmetric measures of the total elasticity for a variable. In the first part of our empirical application, our reported spatial frontier is at the sample mean and is in terms of the mean direct, indirect and total elasticities across the firms to concisely highlight that the fitted model is well-specified. In the second part we use the firm specific parameters from the model to calculate indirect elasticities for individual firms outside the sample mean. These elasticities are then used to obtain the asymmetric bidirectional TFP growth spillovers over the study period, which play a key role in uncovering the spatial TFP growth centers.

The direct, indirect and total elasticities are calculated using the reduced form of our model, which we move towards by rewriting the structural form in Eq. 1 using matrix notation:

$$y_t = \alpha\iota + x_t'\beta + \begin{pmatrix} \mathbf{W_1} x_t'\vartheta_1 + ... \\ +\mathbf{W_M} x_t'\vartheta_M \end{pmatrix} + \begin{pmatrix} \delta_1 \mathbf{W_1} y_t + ... \\ +\delta_M \mathbf{W_M} y_t \end{pmatrix} + \varrho + v_t - \eta - u_t. \quad (6)$$

Here we drop the i and j subscripts from Eq. 1 to denote vectors of stacked cross-sectional observations, ι is an $(N \times 1)$ vector of ones and everything else is as previously defined. $(\delta_1 \mathbf{W_1} y_t + ... + \delta_M \mathbf{W_M} y_t)$ is then taken to the left-hand side giving $(\mathbf{I_N} - \delta_1 \mathbf{W_1} - ... - \delta_M \mathbf{W_M})\, y_t$ and to obtain the following reduced form we divide throughout by $(\mathbf{I_N} - \delta_1 \mathbf{W_1} - ... - \delta_M \mathbf{W_M})$:

$$y_t = \begin{pmatrix} \mathbf{I_N} - \delta_1 \mathbf{W_1} - ... \\ -\delta_M \mathbf{W_M} \end{pmatrix}^{-1} \left(\alpha\iota + x_t'\beta + \begin{pmatrix} \mathbf{W_1} x_t'\vartheta_1 + ... \\ +\mathbf{W_M} x_t'\vartheta_M \end{pmatrix} + \varrho + v_t - \eta - u_t \right). \quad (7)$$

We present the method to calculate the direct and indirect elasticities for the kth exogenous independent variable x_{kt}. As we demonstrate in Eq. 8a, we differentiate Eq. 7 with respect to x_{kt} to obtain a matrix of direct and indirect elasticities for the firms. This matrix of partial derivatives is equal to the matrix product on the right-hand side of Eq. 8b, where this product is independent of the time index. As we noted above, to concisely show that our empirical model is well-specified we present it in terms of mean elasticities across the firms. The mean direct elasticity is the mean of the diagonal elements of the matrix product in Eq. 8b. The mean indirect spillover elasticity can be interpreted as the mean spill-in/spill-out elasticity, which are equal in magnitude, and are the mean row/column sums of the off-diagonal elements of the aforementioned matrix product.[7] Our approach to uncover spatial TFP growth centers, however, uses asymmetric bidirectional indirect elasticities for individual firms. These spill-in and spill-out elasticities are the individual row and column sums of the off-diagonal elements of the matrix product in Eq. 8b.

[7] As we noted in the introductory section, the only mean spillover (i.e., indirect) elasticity from a spatial error/inefficiency model relates to the disturbance/inefficiency. To simplify the discussion we focus on the spatial error model as it is a common specification in the spatial literature. The spatial error model contains no inefficiency component(s) and regarding spatial terms includes only a spatial lag of the disturbance (see the first equation in subsection 5.3 of Viton (2010) on page 11 for the structural form of this model). In contrast to the reduced form of our model in Eq. 7, where the spatial multiplier matrix pre-multiplies all the other components of the model, in the reduced form of the spatial error model the spatial multiplier matrix only pre-multiplies the disturbance (see the final equation in subsection 5.3 of Viton (2010) on page 11). Hence why the only indirect elasticity from this reduced form relates to the disturbance.

$$
\left(\frac{\partial y}{\partial x_{k,1}}, \quad \cdots \quad , \frac{\partial y}{\partial x_{k,N}} \right)_t = \begin{bmatrix} \frac{\partial y_1}{\partial x_{k,1}} & \cdots & \frac{\partial y_1}{\partial x_{k,N}} \\ \vdots & \ddots & \vdots \\ \frac{\partial y_N}{\partial x_{k,1}} & \cdots & \frac{\partial y_N}{\partial x_{k,N}} \end{bmatrix}_t \tag{8a}
$$

$$
= \left(\mathbf{I_N} - \delta_1 \mathbf{W_1} - ... - \delta_M \mathbf{W_M}\right)^{-1} \times
\begin{bmatrix} \beta_k & \cdots & \begin{matrix} w_{1N,1}\vartheta_{1,k} + ... \\ +w_{1N,M}\vartheta_{M,k} \end{matrix} \\ \vdots & \ddots & \vdots \\ \begin{matrix} w_{N1,1}\vartheta_{1,k} + ... \\ +w_{N1,M}\vartheta_{M,k} \end{matrix} & \cdots & \beta_k \end{bmatrix}. \tag{8b}
$$

Following the convention in the spatial econometrics literature (e.g., LSP), we conduct statistical inference for the mean direct, indirect and total elasticities using the standard deviation of $1,000$ Monte Carlo simulations of each mean elasticity. Each simulation is based on drawing a set of parameter values from the variance-covariance estimates that form the Hessian matrix, where each parameter value has a random component drawn from $N\left(0,1\right)$. Alternative approaches can be used to conduct statistical inference by using sequences of a quasi-random nature such as Halton, Faure or Sobol sequences. Using Halton sequences is appealing for reasons of conceptual simplicity and computational speed (Bhat, 2001). For mixed logit models Bhat observes that using 75 Halton draws for statistical inference is more accurate than using $2,000$ pseudo-random draws. To the best of our knowledge, however, there is no corresponding study for any spatial econometric models and, as a result, this is an important area for future research. Such research would provide the foundation for the use of alternative approaches to statistical inference in the spatial literature going forward.

Based on how we calculate the direct and asymmetric bidirectional indirect and total elasticities, we calculate the corresponding efficiency measures for each firm, which collectively we refer to as spatial efficiencies because each of these efficiencies is partially/entirely made up of a spillover. Direct efficiency is interpreted in the same way as own efficiency from a non-spatial or spatial frontier, but also includes a further component as it is own efficiency plus efficiency feedback. Efficiency feedback occurs when, for example, a change in a firm's dependent variable due to a change in one of its independent variables affects the dependent variables of the other firms in the sample and thus their efficiencies. Via the spatial multiplier matrix this effect partially rebounds to the dependent variable and efficiency of the firm that initiated the process. In line with the indirect elasticities for individual firms we calculate asymmetric bidirectional indirect efficiencies, which are the sum of the efficiencies that spill-in (spill-out) to (from) a firm from (to) all the other firms in the sample. We can therefore interpret the spill-in efficiency as being the combined effect of the M networks on an individual firm's efficiency, and the spill-out efficiency as being the contribution of an individual firm to the efficiency of the M networks as a whole. Mirroring the situation for the asymmetric indirect and total elasticities for individual firms we obtain asymmetric total efficiency measures.

The direct and asymmetric bidirectional indirect and total efficiencies we calculate are all absolute measures. Glass *et al.* (2016), on the other hand, calculate relative spatial efficiencies. Relative efficiencies can be sensitive to the best performing firm in each period being an outlier, which one would need to adjust for. Additionally, relative indirect spill-in and spill-out efficiencies do not indicate whether the absolute magnitude of the efficiency spillover is substantive. As a result, there can be cases where it is of no consequence whether the relative indirect efficiencies are high or low because the absolute indirect efficiencies used to calculate these relative measures are small. We therefore calculate absolute measures of the direct and asymmetric bidirectional indirect and total efficiencies by adapting the method in G&K for a single spatial network to our more complex setting of multiple networks.

Using the reduced form of our model in Eq. 7, we can relate the own NIE, NVE and GVE measures to the absolute measures of the direct and asymmetric bidirectional indirect and total efficiencies. This involves recognizing from Eq. 7 that $(\mathbf{I_N} - \delta_1\mathbf{W_1} - ... - \delta_M\mathbf{W_M})^{-1}\eta = \eta_{In}^{Tot}$, $(\mathbf{I_N} - \delta_1\mathbf{W_1} - ... - \delta_M\mathbf{W_M})^{-1}u_t = u_{In,t}^{Tot}$ and $(\mathbf{I_N} - \delta_1\mathbf{W_1} - ... - \delta_M\mathbf{W_M})^{-1}(\eta + u_t) = \eta_{In}^{Tot} + u_{In,t}^{Tot}$ are $(N \times 1)$ vectors of absolute total (Tot) net time-invariant, net time-varying and gross time-varying inefficiencies, respectively. The subscript In attached to these total inefficiency vectors indicates that the inefficiency spillovers used in the calculation of the total inefficiencies are the inefficiency spill-ins to the ith firm from all of the jth firms where $i \neq j$. Since from the structural form of our model in Eq. 1 own $NIE_i = \exp(-\eta_i)$ and own $NVE_{it} = \exp(-u_{it})$, the $(N \times 1)$ vectors of absolute total net efficiencies that directly correspond to the absolute total net inefficiency vectors, η_{In}^{Tot} and $u_{In,t}^{Tot}$, are $NIE_{In}^{Tot} = (\mathbf{I_N} - \delta_1\mathbf{W_1} - ... - \delta_M\mathbf{W_M})^{-1}\exp(-\eta)$ and $NVE_{In,t}^{Tot} = (\mathbf{I_N} - \delta_1\mathbf{W_1} - ... - \delta_M\mathbf{W_M})^{-1}\exp(-u_t)$. Similarly, as own GVE_{it} is $NIE_i \times NVE_{it} = \exp(-\eta_i - u_{it})$, the $(N \times 1)$ vector of absolute total gross time-varying efficiencies that mirrors the corresponding inefficiency vector, $\eta_{In}^{Tot} + u_{In,t}^{Tot}$, is $GVE_{In,t}^{Tot} = (\mathbf{I_N} - \delta_1\mathbf{W_1} - ... - \delta_M\mathbf{W_M})^{-1} \times \exp(-\eta - u_t)$.

$NVE_{In,t}^{Tot}$ can be represented as follows and using the same form we can also represent NIE_{To}^{Tot} and $GVE_{To,t}^{Tot}$.

$$\begin{pmatrix} \mathbf{I_N} - \delta_1\mathbf{W_1} - \\ ... - \delta_M\mathbf{W_M} \end{pmatrix}^{-1} \begin{pmatrix} NVE_1 \\ \vdots \\ NVE_N \end{pmatrix}_t = \begin{pmatrix} NVE_{11}^{Dir} + & \cdots & +NVE_{1N}^{Ind} \\ \vdots & \ddots & \vdots \\ NVE_{N1}^{Ind} + & \cdots & +NVE_{NN}^{Dir} \end{pmatrix}_t = \begin{pmatrix} NVE_{In,1}^{Tot} \\ \vdots \\ NVE_{In,N}^{Tot} \end{pmatrix}_t, \quad (9)$$

where the NVE_{ijt}^{Dir} component of the $NVE_{In,it}^{Tot}$ element denotes absolute direct NVE for the ith firm. When $i \neq j$ the NVE_{ijt}^{Ind} component of the same element represents the absolute indirect NVE spill-in to the ith firm from the jth firm, and $NVE_{In,it}^{Ind} = \sum_{j=1}^{N} NVE_{ijt}^{Ind}$ is the sum of the absolute indirect NVE spill-ins to the ith firm from all the jth firms. We are therefore additively decomposing the $(N \times 1)$ vector $NVE_{In,t}^{Tot}$ into the $(N \times 1)$ vectors NVE_t^{Dir} and $NVE_{In,t}^{Ind}$. In the same way we can decompose NIE_{In}^{Tot} into NIE^{Dir} and NIE_{In}^{Ind}, and $GVE_{In,t}^{Tot}$ into GVE_t^{Dir} and $GVE_{In,t}^{Ind}$.

Typically $\mathbf{W_1}, ..., \mathbf{W_M}$ will be asymmetric, which means $(\mathbf{I_N} - \delta_1\mathbf{W_1} - ... - \delta_M\mathbf{W_M})^{-1}$ will also be asymmetric. As a result, there will be asymmetric spill-ins and spill-outs of NVE, NIE and GVE. G&K set out the methodology for NVE, NIE and GVE spill-outs, although they did not calculate these measures in their empirical application, whereas in our application we do so to provide an empirical demonstration of these spill-outs. Recall that an element of $NVE_{In,t}^{Tot}$ is the horizontal sum of the components in Eq. 9, whereas an element of $NVE_{Out,t}^{Tot}$ is the vertical sum of these components. The subscript Out denotes that the total efficiency vector is calculated using the sum of the absolute indirect NVE spill-outs from all the ith firms to a jth firm ($NVE_{Out,jt}^{Ind} = \sum_{i=1}^{N} NVE_{ijt}^{Ind}$). We are therefore additively decomposing the $(1 \times N)$ vector NVE_{Out}^{Tot} into the $(1 \times N)$ vectors NVE_t^{Dir} and $NVE_{Out,t}^{Ind}$. In the same way we decompose the NIE_{Out}^{Tot} and $GVE_{Out,t}^{Tot}$ vectors.

Own NIE, NVE and GVE are standard measures of efficiency from a stochastic frontier model and are therefore bounded in the interval $[0, 1]$. Direct, asymmetric bidirectional indirect and total NIE, NVE and GVE also have a lower bound of 0. All these spatial efficiency measures, however, have no upper bound, but this in no way prevents the changes in these efficiencies from being included in our spatial TFP growth decompositions. This is because these spatial efficiency measures are scaled own NIE, NVE and GVE. Consequently, if the relevant efficiency spill-in/spill-out, which represents an efficiency performance multiplier, is sufficiently large, the direct/indirect/total NIE, NVE or GVE measure will be greater than 1. In this situation the efficiency spill-in/spill-out is large enough to place the firm beyond the corresponding own best practice frontier from Eq. 1.

3 Methodology to Uncover Spatial TFP Growth Centers

Our methodology to uncover firms that act as spatial TFP growth centers comprises three parts. The first part relates to the specification of the technology that is modeled by the spatial Durbin frontier, although our approach to uncover centers is in no way limited to a particular technology (cost, revenue, etc.). The second part shows how the frontier model can be used to obtain asymmetric bidirectional decompositions of two measures of spatial TFP growth. The two decompositions differ depending on whether the spillover components represent spill-ins to a firm or spill-outs from a firm. In the third part, the TFP growth spill-ins and spill-outs are combined to develop a method to uncover which firms act as spatial TFP growth centers.

3.1 Spatial Revenue Frontier with Multiple Spatial Networks

We complement the G&K spatial cost frontier based application of their spatial TFP growth decomposition by focusing in our application on the income side of a firm's operations. We estimate a spatial Durbin stochastic revenue frontier, which we present by tailoring the general presentation of our model in Eq. 1 to the specific functional form we employ, where unless otherwise stated everything is as defined for Eq. 1.

$$r_{it} = \alpha + TL\left(q_{it}, p_{it}, t_i\right) + \begin{pmatrix} STL_1\left(\sum_{j=1}^{N} w_{ij1}\left(q_{jt}\right), \sum_{j=1}^{N} w_{ij1}\left(p_{jt}\right), \sum_{j=1}^{N} w_{ij1}\left(t_j\right)\right) + ... \\ +STL_M\left(\sum_{j=1}^{N} w_{ijM}\left(q_{jt}\right), \sum_{j=1}^{N} w_{ijM}\left(p_{jt}\right), \sum_{j=1}^{N} w_{ijM}\left(t_j\right)\right) \end{pmatrix} + \begin{pmatrix} \delta_1 \sum_{j=1}^{N} w_{ij1} r_{jt} + ... \\ +\delta_M \sum_{j=1}^{N} w_{ijM} r_{jt} \end{pmatrix} + \varrho_i + v_{it} - \eta_i - u_{it}, \qquad (10)$$

where all the variables are logged. r_{it} is a total revenue observation, and $TL\left(q_{it}, p_{it}, t_i\right) = \omega t_i + \frac{1}{2}\varsigma t_i^2 + \theta' q_{it} + \psi' p_{it} + \frac{1}{2} p_{it}' \boldsymbol{\Upsilon} p_{it} + \frac{1}{2} y_{it}' \boldsymbol{\Omega} y_{it} + p_{it}' \boldsymbol{\Psi} y_{it} + \gamma' q_{it} t + \varphi' p_{it} t$ and represents the variable returns to scale translog approximation of the log of the revenue frontier technology. q_{it} is the $(1 \times H)$ vector of observations for the inputs which are indexed $h \in 1, ..., H$; p_{it} is the $(1 \times L)$ vector of observations for the normalized output prices which are indexed $l \in 1, ..., L$; and t, t^2 and the interactions with t collectively represent a non-linear time trend that measures non-neutral technological change. $STL_1 + ... + STL_M$ represent M spatial lags of $TL\left(q_{it}, p_{it}, t_i\right)$, where these lags together with $\delta_1 \sum_{j=1}^{N} w_{ij1} r_{jt} + ... + \delta_M \sum_{j=1}^{N} w_{ijM} r_{jt}$ shift the frontier technology.[8] To be estimated, among other things, are various regression parameters (α, ω, $\frac{1}{2}\varsigma$, ω_{sm} and $\frac{1}{2}\varsigma_{sm}$, where the sm subscript denotes a local spatial parameter for the mth network); vectors of parameters (θ', ψ', γ', φ', θ'_{sm}, ψ'_{sm}, γ'_{sm} and φ'_{sm}); and matrices of parameters ($\frac{1}{2}\boldsymbol{\Upsilon}$, $\frac{1}{2}\boldsymbol{\Omega}$, $\boldsymbol{\Psi}$, $\frac{1}{2}\boldsymbol{\Upsilon_{sm}}$, $\frac{1}{2}\boldsymbol{\Omega_{sm}}$ and $\boldsymbol{\Psi_{sm}}$). From the properties of the translog function (Christensen *et al.*, 1973) Eq.

[8] In the functions denoted $STL_1, ..., STL_M$ in Eq. 10 the variables are in brackets to indicate that in addition to spatially lagging the variables, we also spatially lag functions of the variables (e.g., q_{jt}^2).

10 is twice differentiable with respect to a normalized output price, an input and the M spatial lags of the normalized output prices and inputs. The resulting Hessian is symmetric due to the symmetry restrictions placed on the parameter matrices.

3.2 Asymmetric Bidirectional Spatial TFP Growth Decompositions

In practice the approach to parametric non-spatial TFP growth decompositions involves first calculating the components, which are then summed to obtain TFP growth. We also adopt this approach here, which involves for the two spatial TFP growth decompositions we employ summing the direct and indirect components to obtain two measures of total TFP growth. Our first decomposition corresponds with that in G&K as it consists of direct and indirect spill-in components to a firm, which gives the same type of total TFP growth measure. Here we use a more sophisticated spatial frontier with multiple networks, and, as a result, total TFP growth and its direct and indirect components will differ in magnitude from what we would obtain using the simpler spatial frontier with a single network in G&K. Despite our first and second decompositions having the same direct components, the decompositions differ because the indirect components in the second formulation gravitate in the opposite direction as they are spill-outs. As a result, we obtain asymmetric bidirectional measures of total TFP growth from the two decompositions.

Our first decomposition is based on the direct and indirect spill-in translog revenue models and the resulting total model. The forms of these models are given in Eqs. $11-13$. These models are constructed by incorporating the direct, indirect spill-in and total parameters (from the matrix product in Eq. 8b) into the reduced form of our model in Eq. 7.[9] For our second decomposition we again use the direct model and also the indirect spill-out model in Eq. 14 and the associated total model in Eq. 15.

$$
\begin{aligned}
r_{it}^{Dir} = {} & \omega_i^{Dir} t_i + \frac{1}{2}\varsigma_i^{Dir} t_i^2 + \theta_i^{Dir'} q_{it} + \psi_i^{Dir'} p_{it} + \frac{1}{2} q_{it}' \mathbf{\Upsilon_i^{Dir}} q_{it} + \frac{1}{2} p_{it}' \mathbf{\Omega_i^{Dir}} p_{it} + \\
& q_{it}' \mathbf{\Psi_i^{Dir}} p_{it} + \gamma_i^{Dir'} q_{it} t_i + \varphi_i^{Dir'} p_{it} t_i - \eta_i^{Dir} - u_{it}^{Dir},
\end{aligned}
\tag{11}
$$

$$
\begin{aligned}
r_{In,it}^{Ind} = {} & \omega_{In,i}^{Ind} t_i + \frac{1}{2}\varsigma_{In,i}^{Ind} t_i^2 + \theta_{In,i}^{Ind'} q_{it} + \psi_{In,i}^{Ind'} p_{it} + \frac{1}{2} q_{it}' \mathbf{\Upsilon_{In,i}^{Ind}} q_{it} + \frac{1}{2} p_{it}' \mathbf{\Omega_{In,i}^{Ind}} p_{it} + \\
& q_{it}' \mathbf{\Psi_{In,i}^{Ind}} p_{it} + \gamma_{In,i}^{Ind'} q_{it} t_i + \varphi_{In,i}^{Ind'} p_{it} t_i - \eta_{In,i}^{Ind} - u_{In,it}^{Ind},
\end{aligned}
\tag{12}
$$

$$
\begin{aligned}
r_{In,it}^{Tot} = {} & \omega_{In,i}^{Tot} t_i + \frac{1}{2}\varsigma_{In,i}^{Tot} t_i^2 + \theta_{In,i}^{Tot'} q_{it} + \psi_{In,i}^{Tot'} p_{it} + \frac{1}{2} q_{it}' \mathbf{\Upsilon_{In,i}^{Tot}} q_{it} + \frac{1}{2} p_{it}' \mathbf{\Omega_{In,i}^{Tot}} p_{it} + \\
& q_{it}' \mathbf{\Psi_{In,i}^{Tot}} p_{it} + \gamma_{In,i}^{Tot'} q_{it} t_i + \varphi_{In,i}^{Tot'} p_{it} t_i - \eta_{In,i}^{Tot} - u_{In,it}^{Tot},
\end{aligned}
\tag{13}
$$

$$
\begin{aligned}
r_{Out,jt}^{Ind} = {} & \omega_{Out,j}^{Ind} t_i + \frac{1}{2}\varsigma_{Out,j}^{Ind} t_i^2 + \theta_{Out,j}^{Ind'} q_{it} + \psi_{Out,j}^{Ind'} p_{it} + \frac{1}{2} q_{it}' \mathbf{\Upsilon_{Out,j}^{Ind}} q_{it} + \frac{1}{2} p_{it}' \mathbf{\Omega_{Out,j}^{Ind}} p_{it} + \\
& q_{it}' \mathbf{\Psi_{Out,j}^{Ind}} p_{it} + \gamma_{Out,j}^{Ind'} q_{it} t_i + \varphi_{Out,j}^{Ind'} p_{it} t_i - \eta_{Out,j}^{Ind} - u_{Out,jt}^{Ind},
\end{aligned}
\tag{14}
$$

[9] This is with the caveat that there are no direct, indirect and total intercepts in Eqs. $11-13$, which is also the case for the other models we utilize for our second spatial TFP growth decomposition. We could obtain these intercepts using a similar approach to how we calculate the corresponding revenue inefficiencies, but this would be an unnecessary complication. This is because we are interested in the factors that lead to changes in the various measures of revenue (direct revenue, indirect revenue spill-ins and spill-outs, and total revenue), whereas the associated intercepts are fixed parameters.

$$r^{Tot}_{Out,jt} = \omega^{Tot}_{Out,j}t_i + \frac{1}{2}\varsigma^{Tot}_{Out,j}t_i^2 + \theta^{Tot'}_{Out,j}q_{it} + \psi^{Tot'}_{Out,j}p_{it} + \frac{1}{2}q'_{it}\mathbf{\Upsilon^{Tot}_{Out,j}}q_{it} + \frac{1}{2}p'_{it}\mathbf{\Omega^{Tot}_{Out,j}}p_{it} + q'_{it}\mathbf{\Psi^{Tot}_{Out,j}}p_{it} + \gamma^{Tot'}_{Out,j}q_{it}t_i + \varphi^{Tot'}_{Out,j}p_{it}t_i - \eta^{Tot}_{Out,j} - u^{Tot}_{Out,jt}, \quad (15)$$

An indirect spill-in elasticity measures the responsiveness of the revenue that spills over to the ith firm, $r^{Ind}_{In,it}$, due to the spill-in effect on an ith firm's independent variable. An indirect spill-out elasticity, on the other hand, measures the responsiveness of the revenue that spills over to the jth firm, $r^{Ind}_{Out,jt}$, due to the spill-out effect from an ith firm's independent variable. Each total parameter is obtained by summing the direct and indirect spill-in/spill-out parameters. Similarly, a firm's asymmetric total inefficiencies are the sum of its direct inefficiency and its indirect inefficiency spill-ins/spill-outs. Inefficiency spill-ins/spill-outs have negative connotations, whereas in practice firms seek to exploit agglomeration economies, which are the benefits from locating near one another in cities and industrial clusters. It is therefore logical to transform the inefficiency spill-ins and spill-outs into efficiency measures using the method in subsection 2.3, which is the approach we adopt in our spatial TFP growth decompositions.

We do not observe r^{Dir}_{it}, $r^{Ind}_{In,it}$, $r^{Tot}_{In,it}$, $r^{Ind}_{Out,jt}$ and $r^{Tot}_{Out,jt}$. They can though, if necessary, be computed using the above translog models, where $r^{Tot}_{In,it} = r^{Dir}_{it} + r^{Ind}_{In,it}$ and $r^{Tot}_{Out,jt} = r^{Dir}_{it} + r^{Ind}_{Out,jt}$ because for r^{Dir}_{it} $i = j$. Note also that each observation in the direct, indirect and total translog models is not pre-multiplied by the sum of the spatial weights. This is because the effect of the spatial weights is incorporated within the direct, indirect and total parameters.

From our two decompositions we obtain the asymmetric bidirectional generalized Malmquist total TFP growth indices ($\Delta TFP^{Tot}_{In,it+1}$ and $\Delta TFP^{Tot}_{Out,jt+1}$) by summing their components. We obtain our two decompositions by following the corresponding non-spatial (i.e., own) TFP growth decomposition (Orea, 2002). Accordingly, our first decomposition in Eq. 16 is the result of fully differentiating the direct and indirect spill-in translog functions in Eqs. 11 and 12.[10] By replacing the full derivative of Eq. 12 in the first decomposition with the full derivative of the indirect spill-out translog function in Eq. 14, we obtain our second decomposition in Eq. 17.

$$\begin{aligned}\Delta TFP^{Tot}_{In,it+1} &= \Delta TC^{Tot}_{In,it+1} + \Delta RTS^{Tot}_{In,it+1} + \Delta GVE^{Tot}_{In,it+1} \\ &= \Delta TFP^{Dir}_{it+1} + \Delta TFP^{Ind}_{In,it+1} \\ &= \Delta TC^{Dir}_{it+1} + \Delta RTS^{Dir}_{it+1} + \Delta GVE^{Dir}_{it+1} + \Delta TC^{Ind}_{In,it+1} + \Delta RTS^{Ind}_{In,it+1} + \\ &\Delta GVE^{Ind}_{In,it+1}.\end{aligned} \quad (16)$$

$$\begin{aligned}\Delta TFP^{Tot}_{Out,jt+1} &= \Delta TC^{Tot}_{Out,jt+1} + \Delta RTS^{Tot}_{Out,jt+1} + \Delta GVE^{Tot}_{Out,jt+1} \\ &= \Delta TFP^{Dir}_{jt+1} + \Delta TFP^{Ind}_{Out,jt+1} \\ &= \Delta TC^{Dir}_{jt+1} + \Delta RTS^{Dir}_{jt+1} + \Delta GVE^{Dir}_{jt+1} + \Delta TC^{Ind}_{Out,jt+1} + \Delta RTS^{Ind}_{Out,jt+1} + \\ &\Delta GVE^{Ind}_{Out,jt+1}.\end{aligned} \quad (17)$$

The components of our two spatial TFP growth decompositions are as follows.

1. $\Delta TC^{Tot}_{In,it+1}$ and $\Delta TC^{Tot}_{Out,jt+1}$ are the total technological change components for the ith and jth firms, respectively. Each of these components relates to an upward or downward shift in the

[10] There is slight abuse of terminology here because by full differentiation we mean total differentiation. To avoid confusion with the total terminology we use from the spatial literature, we instead use full as the descriptor.

relevant total frontier for a firm when there is total technological progress or regress for the firm. To decompose $\Delta TC^{Tot}_{In,it+1}$ and $\Delta TC^{Tot}_{Out,jt+1}$ in Eqs. 18 and 19 into their direct and indirect technological change components, we use the following first order derivatives of the direct and indirect spill-in and spill-out translog functions with respect to time.

$$\begin{aligned}\Delta TC^{Tot}_{In,it+1} &= \frac{1}{2}\left(\frac{\partial r^{Tot}_{In,it+1}}{\partial t_i} + \frac{\partial r^{Tot}_{In,it}}{\partial t_i}\right) \\ &= \frac{1}{2}\left[\left(\frac{\partial r^{Dir}_{it+1}}{\partial t_i} + \frac{\partial r^{Ind}_{In,it+1}}{\partial t_i}\right) + \left(\frac{\partial r^{Dir}_{it}}{\partial t_i} + \frac{\partial r^{Ind}_{In,it}}{\partial t_i}\right)\right] .\end{aligned} \tag{18}$$

$$\begin{aligned}\Delta TC^{Tot}_{Out,jt+1} &= \frac{1}{2}\left(\frac{\partial r^{Tot}_{Out,jt+1}}{\partial t_i} + \frac{\partial r^{Tot}_{Out,jt}}{\partial t_i}\right) \\ &= \frac{1}{2}\left[\left(\frac{\partial r^{Dir}_{jt+1}}{\partial t_i} + \frac{\partial r^{Ind}_{Out,jt+1}}{\partial t_i}\right) + \left(\frac{\partial r^{Dir}_{jt}}{\partial t_i} + \frac{\partial r^{Ind}_{Out,jt}}{\partial t_i}\right)\right] .\end{aligned} \tag{19}$$

2. $\Delta GVE^{Tot}_{In,it+1}$ and $\Delta GVE^{Tot}_{Out,jt+1}$ measure the rise or fall in the total gross time-varying efficiency of the *ith* and *jth* firms. To decompose $\Delta GVE^{Tot}_{In,it+1}$ and $\Delta GVE^{Tot}_{Out,jt+1}$ into their direct and indirect components we use Eqs 20 and 21.

$$\begin{aligned}\Delta GVE^{Tot}_{In,it+1} &= GVE^{Tot}_{In,it+1} - GVE^{Tot}_{In,it} \\ &= \left(GVE^{Dir}_{it+1} + GVE^{Ind}_{In,it+1}\right) - \left(GVE^{Dir}_{it} + GVE^{Ind}_{In,it}\right) .\end{aligned} \tag{20}$$

$$\begin{aligned}\Delta GVE^{Tot}_{Out,jt+1} &= GVE^{Tot}_{Out,jt+1} - GVE^{Tot}_{Out,jt} \\ &= \left(GVE^{Dir}_{jt+1} + GVE^{Ind}_{Out,jt+1}\right) - \left(GVE^{Dir}_{jt} + GVE^{Ind}_{Out,jt}\right) .\end{aligned} \tag{21}$$

3. $\Delta RTS^{Tot}_{In,it+1}$ and $\Delta RTS^{Tot}_{Out,jt+1}$ are the changes in the total returns to scale for the *ith* and *jth* firms. To decompose $\Delta RTS^{Tot}_{In,it+1}$ and $\Delta RTS^{Tot}_{Out,jt+1}$ into their direct and indirect components we use Eqs. 22 and 23.

$$\begin{aligned}\Delta RTS^{Tot}_{In,it+1} &= \frac{1}{2}\left[\sum_{h=1}^{H}\left(e^{Tot}_{h,In,it+1}SF^{Tot}_{In,it+1} + e^{Tot}_{h,In,it}SF^{Tot}_{In,it}\right)\ln\left(\frac{q_{h,it+1}}{q_{h,it}}\right)\right] \\ &= \frac{1}{2}\left[\sum_{h=1}^{H}\left(e^{Dir}_{h,it+1}SF^{Tot}_{In,it+1} + e^{Dir}_{h,it}SF^{Tot}_{In,it}\right)\ln\left(\frac{q_{h,it+1}}{q_{h,it}}\right)\right] + \\ &\quad \frac{1}{2}\left[\sum_{h=1}^{H}\left(e^{Ind}_{h,In,it+1}SF^{Tot}_{In,it+1} + e^{Ind}_{h,In,it}SF^{Tot}_{In,it}\right)\ln\left(\frac{q_{h,it+1}}{q_{h,it}}\right)\right] ,\end{aligned} \tag{22}$$

$$\Delta RTS^{Tot}_{Out,jt+1} = \frac{1}{2}\left[\sum_{h=1}^{H}\left(e^{Tot}_{h,Out,jt+1}SF^{Tot}_{Out,jt+1} + e^{Tot}_{h,Out,jt}SF^{Tot}_{Out,jt}\right)\ln\left(\frac{q_{h,it+1}}{q_{h,it}}\right)\right]$$
$$= \frac{1}{2}\left[\sum_{h=1}^{H}\left(e^{Dir}_{h,jt+1}SF^{Tot}_{Out,jt+1} + e^{Dir}_{h,jt}SF^{Tot}_{Out,jt}\right)\ln\left(\frac{q_{h,it+1}}{q_{h,it}}\right)\right] +$$
$$\frac{1}{2}\left[\sum_{h=1}^{H}\left(e^{Ind}_{h,Out,jt+1}SF^{Tot}_{Out,jt+1} + e^{Ind}_{h,Out,jt}SF^{Tot}_{Out,jt}\right)\ln\left(\frac{q_{h,it+1}}{q_{h,it}}\right)\right]. \quad (23)$$

e denotes a revenue elasticity with respect to the hth input of the ith firm, where these elasticities are obtained from the direct, indirect and total translog functions. SF^{Tot} denotes a total scale factor and to illustrate for our first decomposition $SF^{Tot}_{In,it+1} = \left(\sum_{h=1}^{H} e^{Tot}_{In,h,it+1} + 1\right) / \sum_{h=1}^{H} e^{Tot}_{In,h,it+1}$. To additively decompose $\Delta RTS^{Tot}_{In,it+1}$ and $\Delta RTS^{Tot}_{Out,it+1}$ into changes in direct and indirect spill-in and spill-out returns to scale, we weight the contributions of the direct and indirect spill-in and spill-out input elasticities by the relevant total scale factor.

3.3 Combining Bidirectional TFP Growth Spillovers to Uncover the Centers

We define a firm as being a spatial TFP growth center if it is a net generator of TFP growth spillovers. This is the case if $\Delta TFP^{Ind}_{Out,jt+1} - \Delta TFP^{Ind}_{In,it+1} > 0$, where $\Delta TFP^{Ind}_{Out,jt+1} = \Delta TC^{Ind}_{Out,jt+1} + \Delta RTS^{Ind}_{Out,jt+1} + \Delta GVE^{Ind}_{Out,jt+1}$ and $\Delta TFP^{Ind}_{In,it+1} = \Delta TC^{Ind}_{In,it+1} + \Delta RTS^{Ind}_{In,it+1} + \Delta GVE^{Ind}_{In,it+1}$. We adopt this definition for a spatial TFP growth center because, other things being equal, firms would wish to locate near a firm that is a net generator of TFP growth spillovers. For other firms $\Delta TFP^{Ind}_{Out,jt+1} - \Delta TFP^{Ind}_{In,it+1} < 0$ as they are net recipients of TFP growth spillovers. Typically, the sample of firms is large in an empirical application so there will be a large number of net generators and net recipients of spatial TFP growth spillovers. The interesting issue therefore is which of these firms are the biggest net generators and which are the biggest net recipients.

Using the components of $\Delta TFP^{Ind}_{Out,jt+1}$ and $\Delta TFP^{Ind}_{In,it+1}$ we uncover the key factors that determine which firms act as spatial TFP growth centers. We use these components to establish if a firm is a net generator of spillovers of: (i) growth in technological change; (ii) returns to scale change; and (iii) gross time-varying efficiency change. To illustrate, a firm is a net generator of returns to scale change spillovers if $\Delta RTS^{Ind}_{Out,jt+1} - \Delta RTS^{Ind}_{In,it+1} > 0$.

4 Empirical Application to U.S. Banks

4.1 Data and Specification of the Linkages in the Spatial Networks

The data set is a rich balanced panel of annual observations covering the period $1998 - 2015$ for 387 medium-sized and large commercial U.S. banks.[11] This period is interesting because it includes very different bank operating environments. To investigate the impact of the different operating environments on which banks are the largest net generators of TFP growth spillovers, we use the parameters from our fitted translog spatial revenue frontier to calculate elasticities outside the sample mean. Using these elasticities we obtain our new annual spatial TFP growth decomposition. This annual decomposition then forms part of the analysis to: (i) uncover which banks act as spatial TFP growth centers in 2015, as

[11] We construct our data sample by following Berger and Roman (2017) by classifying a U.S. bank with total assets between \$1 billion and \$3 billion in 2015 as medium-sized, and a bank with total assets greater than \$3 billion in 2015 as large. Based on this classification both bank size categories are well represented in our data sample with 218 medium-sized banks and 169 large.

the final year of our analysis is the most relevant for future policy making; and (ii) investigate if there are differences between these results for 2015 and those for two subperiods- pre-crisis ($1998-2007$) and crisis and beyond ($2008-2015$). According to Berger and Bouwman (2013), the timeline of the U.S. subprime lending crisis spanned the period 2007:Q3 – 2009:Q4. As they regard the majority of 2007 as being pre-crisis and given we use annual data, we take 2007 as being the last year of the pre-crisis subperiod. Moreover, we use balanced panel data (i.e., we consider the continuously operating medium-sized and large banks over our study period) to make like-for-like comparisons between spatial TFP growth in different years.

To specify the two spatial weights matrices in our spatial Durbin stochastic revenue frontier we use the linkages between banks across two branch type networks- brick and mortar branches, BM, and all other types of branches, O. We discuss in detail how we specify the spatial weights further in this section, but in essence we use disaggregated information on the number of bank branches and their locations to calculate the state level geographical overlap of banks' branch networks. We therefore confine our analysis to medium-sized and large banks, as their sufficiently extensive branch networks leads to a sufficient amount of network overlap. As a result, there is not a lack of interconnectedness between the banks in our spatial weights matrices, which would be the case if we also included the large number of small U.S. banks.

We explored using an alternative study period and thus an alternative pre-crisis subperiod that go back to 1994. This was for two reasons. First, the state branch location information we use to specify the spatial weights matrices, which we will discuss further shortly, is available from 1994. Second, the Interstate Banking and Branching Efficiency (IBBE) Act, otherwise known as the Riegle-Neal Act, came into effect on June 1, 1997, which allowed a bank to open branches outside its state of origin.[12] This resulted in greater opportunities for banks to have overlapping branch networks leading to greater potential spatial dependence between banks. Many states though implemented this Act in advance of the effective date. See table 1 in Dick (2006) for the date when each state implemented the Act. For example, Oregon was the first contiguous state to do so on 2/27/95, while twelve states were the last to implement it on 6/1/97.[13]

Compared to a study period of $1998-2015$, when we experimented using the period $1994-2015$ we find that δ_O is in line with our expectations as it remains positive and significant, while although δ_{BM} remains positive it is not as we would expect because it becomes insignificant. This suggests for $1994-1997$ that there was insufficient positive spatial revenue dependence between banks with overlapping brick and mortar branch networks to yield a significant estimate of δ_{BM} for $1994-2015$. We therefore chose to begin our study period and pre-crisis subperiod in 1998. This is because our results suggest that it has taken some time for banks to exhibit the significant positive spatial revenue dependence we associate with banks that have overlapping multi-state brick and mortar branch networks.

The data for the dependent and independent variables in our spatial Durbin frontier (Eq. 10) is at the level of the banking firm or bank holding company. All of this data was sourced from the Call Reports which are available from the Federal Deposit Insurance Corporation (FDIC). Underlying the specification of the output prices and inputs is the well-established intermediation approach to banking (Sealey and Lindley, 1977). In panel A of table 1 we describe the dependent and independent variables and provide summary statistics for the level data for our study period. In particular, the output prices and inputs we include follows the input-output choices in a leading U.S. banking efficiency paper by

[12] All but two states committed to the nationwide branching in the Act on or prior to the 6/1/97 federal deadline. The two states that opted out ahead of this deadline were Montana and Texas.

[13] These twelve states are as follows: Colorado; Georgia; Hawaii; Illinois; Kansas; Kentucky; Louisiana; Minnesota; Missouri; New Hampshire; Tennessee; and Wisconsin.

Table 1: Description of the data and summary statistics

Panel A: Variable Descriptions	Mean	Std. Dev.
Total revenue (in 000s of 2005 U.S. dollars) (r)	869,838	4,967,991
Price of loans (p_1): Interest income from loans divided by loans and leases	0.0628	0.0173
Price of securities (p_2): Interest income from securities divided by securities	0.0408	0.0267
Price of other activities (p_3): Approximated by total non-interest income divided by total assets	0.0140	0.0317
Fixed assets (in 000s of 2005 U.S. dollars) (q_1): Value of premises and fixed assets	136,177	655,340
Labor (q_2): Number of full-time equivalent employees	2,931	16,070
Total deposits (in 000s of 2005 U.S. dollars) (q_3)	10,953,887	68,987,060
Panel B: Features of the Two Branch Networks	**Mean**	**Std. Dev.**
Number of full service brick and mortar branches	38,659	13,580
Number of other full and limited service branches	4,990	1,643
Panel C: Selected Features of the Other Branch Network	**Mean**	**Std. Dev.**
Number of full service retail offices	3,288	1,198
Number of full service cyber offices	45	21
Number of limited service administrative offices	144	30
Number of limited service military facilities	14	4
Number of limited service mobile/seasonal offices	220	122
Number of limited service trust offices	38	18
Number of limited service drive-through facilities	1,237	285

Koetter *et al.* (2012).

Summarizing, there are three output prices and three inputs in our model. The output prices relate to the prices of the following lending and non-lending activities of banks: loans (p_1), securities (p_2) and non-interest income (p_3). Total revenue (r) is the sum of the revenues from these outputs and the inputs are fixed assets (q_1), labor (q_2) and deposits (q_3), where q_1, q_3 and r are deflated to 2005 prices using the CPI. The output prices are not deflated because as is evident from panel A of table 1 they are ratios. All the variables are then logged, mean adjusted and, finally, we use one of the output prices (p_1) as the normalizing factor for r and the other output prices. By mean adjusting the data all the first order direct, indirect and total input, time and normalized output price parameters can be interpreted as elasticities at the sample mean. This is because in the first order partial derivatives of a translog function the terms relating to the squared and interaction terms are zero at the sample mean.

We explored departing from the input-output choices in Koetter *et al.* by also including equity capital as an input. Although this resulted in a model for our study period where, as we would expect, δ_O remains positive and significant, δ_{BM} is no longer positive and significant and in line with our expectations, and is instead negative and insignificant. An insignificant δ_{BM} suggests that there is no substantive spatial revenue dependence between banks that have overlapping brick and mortar branch networks, which we feel is counterintuitive for banks that operate in the same markets. In line therefore with Koetter *et al.* we omit equity.

For the reasons given in subsection 2.1, our modeling framework is intended for a reasonably small number of fairly distinct simultaneous spatial networks. In practice, many cases, including our application to U.S. banks, can be adapted to this setting by aggregating networks. In particular and as we noted above, we focus on two distinct bank branch networks- the full service brick and mortar branch network and the aggregated network of other types of full service and limited service branches. Other types of branches can potentially cover up to twelve branch types, although there are only seven of the twelve in our sample.[14] We present in panel B of table 1 some summary statistics for the two branch networks we

[14] For the full list of the twelve other branch types see

consider, and in panel C of table 1 we list the seven other branch types that feature in our sample and provide some further summary statistics.

Our two-way split of the bank branch networks is logical for two reasons. First, the split differentiates between the majority of branches in our sample (brick and mortar branches which account for 88.4% of branches) and the relatively small number of other branch types. Second, the split distinguishes between the different degrees of centralization of activities across the two branching networks. To illustrate, across a large number of brick and mortar branches similar general activities are highly decentralized, whereas in some other branch types there is a high degree of centralization of specialist activities. This explains why in panel C of table 1, the sample means of, for example, the number of cyber offices and administrative offices are small. A related issue we examine is whether the degree of SAR revenue dependence across overlapping brick and mortar branch networks is greater than across overlapping networks of other branches due to the dominance of brick and mortar branches in our sample. Alternatively, SAR revenue dependence could be greater across overlapping networks of other branches due to the higher degree of centralization of activities in some of the other branch types.

In our model the two specifications of $\mathbf{W}$ based on brick and mortar branches and other branch types are denoted $\mathbf{W_{BM}}$ and $\mathbf{W_O}$, respectively. Using information from the FDIC in the Summary of Deposits on the state locations of banks' brick and mortar branches and other branch types, in advance of estimating the models we construct $\mathbf{W_{BM}}$ and $\mathbf{W_O}$ in the same way using the following four steps.

(i) Start with a matrix for each year and set all the cells on the main diagonal to zero because a bank cannot be in its own neighborhood set.

(ii) For each state where the *ith* bank has chosen to locate the relevant branch type, we calculate the ratio of the number of *jth* bank branches to the number of *ith* bank branches. This ratio represents the state level branch intensity of the *jth* bank relative to the *ith* bank. The off-diagonal elements are then calculated by summing these ratios across the states. A zero off-diagonal element signifies no overlap between the branch type networks of the *ith* and *jth* banks.[15]

(iii) Average the annual matrices from (ii) to obtain a matrix for the sample.

(iv) Normalize the matrix for the sample by dividing throughout by the largest element (i.e., normalize by the largest eigenvalue). This normalization is appealing because it preserves the proportional relationship between the spatial weights. As a result, the information on the absolute intensity of a bank's branch type network relative to the overlapping network of another bank is retained. This normalization also aids the interpretation of the indirect elasticities. This is because the spillovers to and from a bank will be positively related to the absolute intensities, rather some transformation of these intensities.[16]

As the revenue of a bank in a particular time period depends on, among other things, the number of branches which the bank has and their locations, it is reasonable to conclude for our spatial revenue

https://www5.fdic.gov/sod/definitions.asp?systemform=soddnld3&helpitem=brsertyp&baritem=1.

[15] Using matrix terminology the *ith* bank corresponds to a row of a $\mathbf{W}$ matrix and the *jth* bank relates to a column. Additionally, as other branch types represent a small share of the total number of branches, we calculate the branch intensities for the banks at the rather aggregate state level. This is to ensure there are sufficient overlapping networks of other branch types in our sample.

[16] It is common to normalize the spatial weights in a $\mathbf{W}$ matrix by the row sums. This is appropriate when the spatial weights are binary to reflect, for example, contiguous neighboring regions. This, however, is very different to the non-binary spatial linkages between U.S. banks that we use. If we row-normalized our non-binary spatial weights the information about the absolute branch network intensities would be lost. This is due to the normalizing factor not being the same for all the weights.

frontier model that $\mathbf{W_{BM}}$ and $\mathbf{W_O}$ are exogenously determined. Exogenous multiple spatial weights matrices is an assumption which our modeling framework is based on and is in line with the majority of the spatial literature, although in this literature it is typical to consider the simpler standard case of a single spatial weights matrix.[17]

4.2 Estimated Spatial Revenue Frontier and the Elasticities

The estimated coefficients for our spatial Durbin revenue frontier with common correlated effects ($\mathrm{SDRF_C}$) are presented in table 2.[18] As we have previously noted, the estimated SAR coefficients cannot be interpreted as spillover elasticities. The interpretable spillover parameters from our model are the indirect parameters. In table 3 we present the mean indirect parameters for our sample together with the mean estimates of the direct and total parameters. Note that we do not use these mean parameters to compute the spatial TFP growth decompositions and to uncover which banks act as spatial TFP growth centers, and instead use the direct, indirect and total parameters for the individual banks. This is because mean indirect spill-in and spill-out parameters are equal, whereas these parameters for individual banks differ in magnitude and lead to the asymmetric indirect TFP growth spill-ins and spill-outs we need to uncover which banks act as spatial TFP growth centers. We report the mean direct, indirect and total parameters here, however, to concisely demonstrate that we use a well-specified model for our spatial TFP growth decompositions.

Although a SAR coefficient cannot be interpreted as the elasticity of the spatial lag of the dependent variable or, in other words, a spillover elasticity, it can provide useful information on the degree of SAR dependence across the firms. For models with multiple SAR variables, a SAR parameter can be negative because networks in the model overlap. In this case a network with a negative SAR parameter is offsetting some of the positive SAR dependence across another overlapping network, where a negative SAR parameter is interpreted as evidence of competition between firms in a spatial network (e.g., Kao and Bera, 2013).

From table 2 we can see that the estimates of the SAR coefficients, δ_{BM} (0.167) and δ_O (0.232), are significant at the 5% and 0.1% levels, respectively. In the context of the spatial literature, the magnitudes of these parameters points to non-negligible positive SAR revenue dependence between banks with brick and mortar branches and other types of branches in the same state, which offers support for a spatial approach to revenue modeling for U.S. banks. As δ_{BM} and δ_O are both positive our model indicates that there is not, on average, a sufficient overlap between the two bank branch networks. This suggests that the distribution across the states of a bank's brick and mortar branches is statistically different from the distribution of its other branch types. δ_{BM} though is noticeably less than δ_O, which suggests that the dominance of brick and mortar branches in our sample is a smaller source of SAR revenue dependence than the high degree of centralization of activities in some of the other branch types.

In table 2 a number of local spatial coefficients are significant at the 5% level or less (e.g., the

[17] If there are the necessary methodological advances, an area for further work would be to replace our exogenous multiple spatial weights matrices, which are based on geographical branch information, with endogenous specifications that are constructed using business/economic factors such as branch deposits. The literature on endogenous spatial weights matrices (Qu and Lee, 2015) is in its infancy and considers only a model for cross-sectional data with a single spatial weights matrix, while also assuming that firms are 100% efficient. To fit our framework one would need to extend this model to panel data, multiple spatial weights matrices and inefficient firms. We do not pursue these extensions here as we focus on introducing a methodology for practitioners to uncover spatial TFP growth centers.

[18] With regard to our model specification, we omit spatial lags of t for parsimony as these lags would model the spillover of the same technological change phenomenon that is modeled by t, as t is the same for all the banks. Again for parsimony and to simply illustrate the importance of the mean auxiliary variables, we only include means of the first order inputs and normalized output prices as these are the key technology variables. We omit the mean of t as it would be same for all the banks and would not therefore account for any inter-bank heterogeneity.

Table 2: Estimated spatial revenue frontier with two simultaneous spatial networks

	Model coeff		Model coeff		Model coeff
Constant	0.001	$\mathbf{W_{BM}}q_2$	−0.356*	$\mathbf{W_O}\left(q_1^2\right)$	0.119*
q_1	0.012**	$\mathbf{W_{BM}}q_3$	−0.005	$\mathbf{W_O}\left(q_2^2\right)$	0.125
q_2	0.095***	$\mathbf{W_{BM}}p_2$	0.090	$\mathbf{W_O}\left(q_3^2\right)$	−0.147*
q_3	0.885***	$\mathbf{W_{BM}}p_3$	−0.064	$\mathbf{W_O}\left(q_1q_2\right)$	−0.361**
p_2	0.137***	$\mathbf{W_{BM}}\left(q_1^2\right)$	−0.425***	$\mathbf{W_O}\left(q_1q_3\right)$	0.086
p_3	0.175***	$\mathbf{W_{BM}}\left(q_2^2\right)$	0.114	$\mathbf{W_O}\left(q_2q_3\right)$	0.186
q_1^2	−0.009***	$\mathbf{W_{BM}}\left(q_3^2\right)$	−0.191	$\mathbf{W_O}\left(p_2^2\right)$	−0.049*
q_2^2	0.036***	$\mathbf{W_{BM}}\left(q_1q_2\right)$	0.016	$\mathbf{W_O}\left(p_3^2\right)$	−0.061
q_3	0.030***	$\mathbf{W_{BM}}\left(q_1q_3\right)$	0.733**	$\mathbf{W_O}\left(p_2p_3\right)$	0.120
q_1q_2	−0.002	$\mathbf{W_{BM}}\left(q_2q_3\right)$	−0.288	$\mathbf{W_O}\left(q_1p_2\right)$	−0.127
q_1q_3	0.005	$\mathbf{W_{BM}}\left(p_2^2\right)$	−0.115	$\mathbf{W_O}\left(q_1p_3\right)$	0.198**
q_2q_3	−0.065***	$\mathbf{W_{BM}}\left(p_3^2\right)$	0.064	$\mathbf{W_O}\left(q_2p_2\right)$	−0.168
p_2^2	0.004	$\mathbf{W_{BM}}\left(p_2p_3\right)$	−0.485***	$\mathbf{W_O}\left(q_2p_3\right)$	0.014
p_3^2	0.049***	$\mathbf{W_{BM}}\left(q_1p_2\right)$	0.176	$\mathbf{W_O}\left(q_3p_2\right)$	0.232*
p_2p_3	−0.047***	$\mathbf{W_{BM}}\left(q_1p_3\right)$	−0.468***	$\mathbf{W_O}\left(q_3p_3\right)$	−0.220*
q_1p_2	0.054***	$\mathbf{W_{BM}}\left(q_2p_2\right)$	−0.043	$\mathbf{W_O}\left(q_1t\right)$	0.001
q_1p_3	0.008**	$\mathbf{W_{BM}}\left(q_2p_3\right)$	0.658***	$\mathbf{W_O}\left(q_2t\right)$	−0.013
q_2p_2	−0.030***	$\mathbf{W_{BM}}\left(q_3p_2\right)$	−0.071	$\mathbf{W_O}\left(q_3t\right)$	0.013*
q_2p_3	0.037***	$\mathbf{W_{BM}}\left(q_3p_3\right)$	−0.164	$\mathbf{W_O}\left(p_2t\right)$	−0.006
q_3p_2	−0.020**	$\mathbf{W_{BM}}\left(q_1t\right)$	0.025*	$\mathbf{W_O}\left(p_3t\right)$	0.010*
q_3p_3	−0.034***	$\mathbf{W_{BM}}\left(q_2t\right)$	−0.043**	$\mathbf{W_{BM}}r$	0.167*
t	0.001*	$\mathbf{W_{BM}}\left(q_3t\right)$	0.018	$\mathbf{W_O}r$	0.232***
t^2	−0.001***	$\mathbf{W_{BM}}\left(p_2t\right)$	0.002	τ	0.165***
q_1t	0.003***	$\mathbf{W_{BM}}\left(p_3t\right)$	−0.001	$\overline{q_1}$	−0.104***
q_2t	−0.004***	$\mathbf{W_O}q_1$	−0.014	$\overline{q_2}$	−0.002
q_3t	1×10^{-4}	$\mathbf{W_O}q_2$	−0.224*	$\overline{q_3}$	0.144***
p_2t	−0.006***	$\mathbf{W_O}q_3$	−0.046	$\overline{p_2}$	−0.029***
p_3t	-3×10^{-4}	$\mathbf{W_O}p_2$	0.204**	$\overline{p_3}$	0.004
$\mathbf{W_{BM}}q_1$	0.350***	$\mathbf{W_O}p_3$	0.027		

Note: *, ** and *** denote statistical significance at the 5%, 1% and 0.1% levels, respectively.

coefficients on $\mathbf{W_{BM}}q_1$, $\mathbf{W_{BM}}q_2$, $\mathbf{W_O}q_2$ and $\mathbf{W_O}p_2$). This constitutes support for the spatial Durbin model over the SAR specification as the latter omits local spatial variables. The same table reveals that the coefficients on three of the five mean variables ($\overline{q_1}$, $\overline{q_3}$ and $\overline{p_2}$) are significant at the 0.1% level. Using a χ^2_5 likelihood ratio test of the corresponding random effects model (SDRF_R) against the SDRF_C specification, we reject the null that all the coefficients on the mean variables in the SDRF_C are zero at the 0.1% level. This highlights the importance in our application of accounting for correlation between the firm specific effects and at least the key exogenous regressors.[19]

As we use mean adjusted data all the first order parameters of the direct, indirect and total translog functions in table 3 are elasticities at the sample mean. Since direct parameters can be interpreted in the same way as the standard own parameters from the corresponding non-spatial model, the monotonicity properties of a non-spatial revenue frontier also apply to the direct elasticities from our model. It is evident from table 3 that all the direct first order output price and input parameters are positive. This

[19] Note also that τ in table 2 is the weight attached to the cross-sectional component of the data and is significant at the 0.1% level.

Table 3: Direct, indirect and total parameters

	Direct parameter	Indirect parameter	Total parameter
q_1	0.012**	0.120***	0.132***
q_2	0.095***	−0.111*	−0.016
q_3	0.885***	0.162***	1.047***
p_2	0.138***	0.053*	0.190***
p_3	0.175***	0.009	0.185***
q_1^2	−0.009***	−0.138**	−0.147**
q_2^2	0.036***	−0.081	−0.045
q_3^2	0.030***	−0.075	−0.045
q_1q_2	−0.002	0.116	0.114
q_1q_3	0.006	0.235**	0.241**
q_2q_3	−0.066***	−0.034	−0.100
p_2^2	0.004	0.035	0.038
p_3^2	0.049***	−0.052	−0.003
p_2p_3	−0.048***	−0.351***	−0.399***
q_1p_2	0.054***	0.107	0.162**
q_1p_3	0.008**	−0.037	−0.029
q_2p_2	−0.030**	−0.047	−0.077
q_2p_3	0.037***	0.168**	0.205***
q_3p_2	−0.020**	0.042	0.023
q_3p_3	−0.034***	−0.122*	−0.155**
t	0.001*	1×10^{-4}	0.001*
t^2	−0.001***	-1×10^{-4}***	−0.001***
q_1t	0.003***	0.008*	0.011**
q_2t	−0.004***	−0.021***	−0.025***
q_3t	1×10^{-4}	0.013**	0.013**
p_2t	−0.006***	−0.004	−0.010**
p_3t	-3×10^{-4}	0.005	0.005

Note: *, ** and *** denote statistical significance at the 5%, 1% and 0.1% levels, respectively.

is consistent with production theory as it indicates that our model satisfies the monotonicity properties of the translog revenue frontier at the sample mean. These first order output price and input parameters are significant at the 1% level or less, and the latter point to constant direct returns to scale (0.99). Also from table 3 and in line with our expectations, the first order direct time parameter is positive and significant, which suggests annual technological progress for the sample average bank.

Production theory does not predict the signs of the indirect and total elasticities for the output prices and inputs. Tables 3 reveals that a large number of indirect parameters are significant, which once again justifies our spatial approach to revenue modeling for U.S. banks. To illustrate, the only first order indirect output price and input parameter that is not significant relates to p_3. The indirect q_1, q_3 and p_2 parameters are all positive, whereas the indirect q_2 parameter is negative. The latter is because the negative and significant $\mathbf{W_{BM}}q_2$ and $\mathbf{W_O}q_2$ parameters in table 2 more than offset the positive effect of the two SAR parameters. As a result of the positive and significant direct and indirect q_1, q_3 and p_2 parameters, summing these parameters gives positive and significant total parameters. Interestingly, the negative and significant indirect q_2 parameter offsets its positive and significant direct counterpart, which results in a total parameter that is not significant.

Table 4: Summary of the own, direct, indirect and total efficiencies

	Own NVE	Own NIE	Own GVE	GVE_i^{Dir}	$GVE_{In,i}^{Ind}$	$GVE_{In,i}^{Tot}$	$GVE_{Out,j}^{Ind}$	$GVE_{Out,j}^{Tot}$
Mean	0.938	0.887	0.833	0.833	0.148	0.981	0.148	0.981
Median	0.943	0.902	0.846	0.846	0.075	0.932	0.022	0.881
Std. dev.	0.027	0.065	0.070	0.070	0.209	0.221	0.606	0.615

Note: $GVE_i^{Dir} = GVE_j^{Dir}$

4.3 Estimates of the Different Revenue Efficiency Measures

In table 4 we summarize, among other things, the own NVE, NIE and GVE scores from the fitted structural form of our SDRF_c. Although the structural form of the SDRF_c accounts for global SAR and local spatial interactions, the own efficiencies do not include any form of efficiency spillover. On the other hand, if the relevant type of efficiency spillover is present, the direct, indirect and total efficiencies from the reduced form of the SDRF_c will be partially/entirely made up of a spillover.

From table 4 it is evident that the sample mean own NVE and NIE scores are 0.938 and 0.887. Multiplying these measures yields a sample mean own GVE of 0.833. When we test the fitted structural form SDRF_c for net time-invariant and net time-varying inefficiencies (η and u), we find that η and u are significant as we reject each null at the 1% level ($H_0 : \widehat{\sigma}_\eta^2 = 0$ and $H_0 : \widehat{\sigma}_u^2 = 0$). A feature of the sample mean own NVE, NIE and GVE scores is that they are relatively high, which is intuitive for our sample of U.S. banks because in a number of sub-periods high revenue growth is widespread.

GVE is a more complete representation of economic performance than NIE or NVE, so in table 4 we also summarize the direct, indirect and total GVE results from the reduced form SDRF_c. Although in theory there can be efficiency feedback, this table reveals that the sample mean GVE^{Dir} is 0.833 and hence equal to the sample mean own GVE, which suggests that there is no efficiency feedback. This finding is in line with G&K and Glass *et al.* (2019) who, using similar samples of U.S. banks to our sample, essentially find no GVE feedback from spatial cost and spatial profit stochastic frontiers with a single spatial network.

We can see from table 4 that the sample mean $GVE_{In,i}^{Ind}$ is 0.148 and, as we highlighted in the methodology, is equal to the sample mean $GVE_{Out,j}^{Ind}$. These sample means, although not huge, are non-negligible with $GVE_{In,i}^{Ind}$ and, in particular, $GVE_{Out,j}^{Ind}$ varying a lot across the banks. This is evident from table 4 because although the standard deviation of $GVE_{In,i}^{Ind}$ is large, it is slightly more than a third of the standard deviation of $GVE_{Out,j}^{Ind}$.

In contrast to the symmetric bidirectional sample means of the indirect GVE scores, the GVE spill-ins and spill-outs for individual banks are asymmetric. As a result, these spill-ins and spill-outs for individual banks are more akin to the asymmetric median $GVE_{In,i}^{Ind}$ and $GVE_{Out,j}^{Ind}$ scores in table 4. For $GVE_{In,i}^{Ind}$ and particularly $GVE_{Out,j}^{Ind}$ the median is less than the mean, which is consistent with the distributions being positively skewed. This is intuitive for our sample because a relatively small number of banks with very large $GVE_{In,i}^{Ind}$ and $GVE_{Out,j}^{Ind}$ is consistent with the number of banks that have extensive branch networks, which, as a result, overlap with a lot of other networks. A larger number of banks with small $GVE_{In,i}^{Ind}$ and $GVE_{Out,j}^{Ind}$, on the other hand, is in line with the majority of our sample having much smaller branch networks, which leads to less overlap between the networks.

Note that $GVE_i^{Dir} + GVE_{In,i}^{Ind} = GVE_{In,i}^{Tot}$ and $GVE_j^{Dir} + GVE_{Out,j}^{Ind} = GVE_{Out,j}^{Tot}$. We can see therefore from table 4 that the sample means of $GVE_{In,i}^{Ind}$ and $GVE_{Out,j}^{Ind}$ push the sample mean $GVE_{In,i}^{Tot}$ and $GVE_{Out,j}^{Tot}$ scores close to 1. Given how close the sample means of $GVE_{In,i}^{Tot}$ and $GVE_{Out,j}^{Tot}$ are to 1, unsurprisingly $GVE_{In,i}^{Tot}$ and/or $GVE_{Out,j}^{Tot}$ greater than 1 is the case for a number of individual banks. This indicates for these banks that $GVE_{In,i}^{Ind}$ and/or $GVE_{Out,j}^{Ind}$ have raised performance above the level of the own GVE frontier.

4.4 Spatial TFP Growth Decompositions with Asymmetric Bidirectional Spillovers

In figure 1 we present for the sample three annual average spatial TFP growth decompositions. In this figure we express the spatial TFP growth decompositions as indices as we take the exponential of the growth rates. In all three cases we decompose a spatial measure of TFP growth as each measure and its components are partially/entirely made up of spillovers.

- First, we decompose the ΔTFP_i^{Dir} measure into ΔTC_i^{Dir}, ΔRTS_i^{Dir} and ΔGVE_i^{Dir}. This measure of spatial TFP growth and its components are partially made up of the same type of spillover as they all contain feedback to a bank.

- Second, we decompose the $\Delta TFP_{In,i}^{Ind}$ measure into $\Delta TC_{In,i}^{Ind}$, $\Delta RTS_{In,i}^{Ind}$ and $\Delta GVE_{In,i}^{Ind}$, and the $\Delta TFP_{Out,j}^{Ind}$ measure into $\Delta TC_{Out,j}^{Ind}$, $\Delta RTS_{Out,j}^{Ind}$ and $\Delta GVE_{Out,j}^{Ind}$. The $\Delta TFP_{In,i}^{Ind}$ and $\Delta TFP_{Out,j}^{Ind}$ measures of spatial TFP growth (and their corresponding components) differ as they are entirely made up of spill-ins and spill-outs, respectively. In figure 1, however, we do not distinguish between the inbound and outbound direction of the spillovers and thus between $\Delta TFP_{In,i}^{Ind}$ and $\Delta TFP_{Out,j}^{Ind}$ and each of their corresponding components. This is because in this figure to concisely present the decompositions of $\Delta TFP_{In,i}^{Ind}$ and $\Delta TFP_{Out,j}^{Ind}$, for each year we average across the sample giving $\Delta TFP_{In,i}^{Ind} = \Delta TFP_{Out,j}^{Ind}$, which is also the case for their corresponding components. As we noted previously in the methodological part of the paper, to uncover firms that act as spatial TFP growth centers we cannot use annual symmetric sample average indirect elasticities because they yield $\Delta TFP_{In,i}^{Ind} = \Delta TFP_{Out,i}^{Ind}$. We instead use annual asymmetric indirect elasticities for individual banks, which we will provide evidence of further in this subsection. Using asymmetric indirect elasticities is necessary to yield a group of banks with $\Delta TFP_{Out,j}^{Ind} - \Delta TFP_{In,i}^{Ind} > 0$, which is our criterion to classify a bank as a spatial TFP growth center.

- Third, we decompose the $\Delta TFP_{In,i}^{Tot}$ measure into $\Delta TC_{In,i}^{Tot}$, $\Delta RTS_{In,i}^{Tot}$ and $\Delta GVE_{In,i}^{Tot}$, and the $\Delta TFP_{Out,j}^{Tot}$ measure into $\Delta TC_{Out,j}^{Tot}$, $\Delta RTS_{Out,j}^{Tot}$ and $\Delta GVE_{Out,j}^{Tot}$. Recall that $\Delta TFP_{In,i}^{Tot} = \Delta TFP_i^{Dir} + \Delta TFP_{In,i}^{Ind}$ and $\Delta TFP_{Out,j}^{Tot} = \Delta TFP_j^{Dir} + \Delta TFP_{Out,j}^{Ind}$. Since $\Delta TFP_i^{Dir} = \Delta TFP_j^{Dir}$, any difference between $\Delta TFP_{In,i}^{Tot}$ and $\Delta TFP_{Out,j}^{Tot}$ is entirely due to the difference between $\Delta TFP_{In,i}^{Ind}$ and $\Delta TFP_{Out,j}^{Ind}$. In figure 1 though $\Delta TFP_{In,i}^{Tot} = \Delta TFP_{Out,j}^{Tot}$, which is also the case for their corresponding components. This is because in this figure to concisely present the decompositions of $\Delta TFP_{In,i}^{Tot}$ and $\Delta TFP_{Out,j}^{Tot}$ for each year we average across the sample.

There are four overriding features of figure 1. First, as ΔTFP^{Ind} closely tracks ΔTFP^{Tot} it is clear that it is ΔTFP^{Ind} which is driving ΔTFP^{Tot} and not ΔTFP^{Dir}. Second, there are a number of years in our study period where ΔTFP^{Ind} is substantial, e.g., in $2005-2006$ it is of the order of 8%. Third, in terms of the components of ΔTFP^{Ind}, it is evident that ΔRTS^{Ind} is basically the sole driver of ΔTFP^{Ind} and, as a result, ΔRTS^{Tot} has a similarly important role in the calculation of ΔTFP^{Tot}. Together these findings highlight the importance of analyzing spatial returns to scale in U.S. banking (see Glass *et al.*, 2018b, for such an analysis). Fourth, as the two dark shaded areas in figure 1 represent the two recessions in U.S. economic activity over our study period, we can see that there are clear downturns in ΔTFP^{Ind} during these recessions.[20] There are not similar downturns in ΔTFP^{Dir} in these recessions, which is

[20] Strictly the dark shaded areas represent illustrations of the dates of the two recessions from the National Bureau of Economic Research (NBER). The areas are only illustrative because we use annual data, whereas U.S. recessions are dated by the NBER in months (quarters). The NBER dates of the two recessions in our study period are Mar (Q1) 2001-Nov (Q4) 2001 and Dec (Q3) 2007-June (Q2) 2009.

almost certainly because ΔTFP^{Ind} is more reflective of the economic environment in the U.S. banking industry than ΔTFP^{Dir}. Our reasoning here is that ΔTFP^{Ind} is the result of spillovers between a bank and its industry networks, whereas if we overlook the typically small feedback component, ΔTFP^{Dir} considers a bank net of its spatial interaction with other banks.

The key policy implication from the spatial TFP growth decompositions concerns the impact of the suggestion that size caps should be imposed on very large U.S. banks. The background to this suggestion relates to the key role of the "too-big-to-fail" (TBTF) status of very large U.S. banks in the 2008 financial crisis, as this status promoted excessive risk taking. To guard against a repeat of this risk taking, the 2010 Dodd-Frank reforms involved: tightening the regulatory regime through, for example, more stringent bank liquidity constraints; and establishing a formal process to resolve large bank failures with the intention of ensuring that no bank is TBTF. Fisher and Rosenblum (2012), among others, argue that Dodd-Frank could have gone further to prevent TBTF banks by introducing size caps on the very large banks. From the perspective of a revenue based measure of TFP growth, our results do not support such a policy. Our argument is that if such caps were introduced it would conceivably lead to negative ΔRTS^{Ind} because smaller banks would yield smaller spillovers. Based on our results this will then lead to negative ΔTFP^{Ind} and ΔTFP^{Tot}, as ΔRTS^{Ind} is by far and away the biggest driver of both these measures.

As we noted above, we use annual asymmetric indirect elasticities for individual banks to obtain an annual asymmetry between $\Delta TFP^{Ind}_{Out,j}$ and $\Delta TFP^{Ind}_{In,i}$ for each bank. But we cannot concisely present this annual asymmetry for individual banks for a sufficiently large proportion of the sample. For any subset of the sample, however, the annual average $\Delta TFP^{Ind}_{Out,j}$ and $\Delta TFP^{Ind}_{In,i}$ are asymmetric. We concisely demonstrate this asymmetry in figure 2 by presenting the annual average $\Delta TFP^{Ind}_{Out,j}$ and $\Delta TFP^{Ind}_{In,i}$ for the 1st and 5th quintiles of the $\Delta TFP^{Ind}_{In,i}$ distribution in 2015. All the annual $\Delta TFP^{Ind}_{Out,j}$ and $\Delta TFP^{Ind}_{In,i}$ in figure 2 for each quintile are therefore comparable because they are based on the same subset of banks.

There is a lot of evidence in figure 2 of an annual asymmetry between $\Delta TFP^{Ind}_{Out,j}$ and $\Delta TFP^{Ind}_{In,i}$. This is the case because in the blue shaded areas $\Delta TFP^{Ind}_{Out,j} > \Delta TFP^{Ind}_{In,i}$, which are the periods where the 1st and 5th quintiles are, on average, net generators of TFP growth spillovers. In the red shaded areas $\Delta TFP^{Ind}_{Out,j} < \Delta TFP^{Ind}_{In,i}$, which are the periods where the 1st and 5th quintiles are, on average, net recipients of TFP growth spillovers. Figure 2 also suggests that there are downturns in $\Delta TFP^{Ind}_{In,i}$ for both quintiles around the two recessions. Interestingly, for the 1st quintile there is a downturn in $\Delta TFP^{Ind}_{In,i}$ that precedes and coincides with the second recession, whereas for the 5th quintile there is a downturn in $\Delta TFP^{Ind}_{In,i}$ that immediately follows this recession.

4.5 Spatial TFP Growth Centers and the Drivers

In figure 3 we present the geographical distribution of the branches of the banks that are net generators of TFP growth spillovers in 2015, i.e., $\Delta TFP^{Ind}_{Out,j} - \Delta TFP^{Ind}_{In,i} > 0$. The banks whose branches feature in figure 3 are therefore classified as spatial TFP growth centers. In this figure we are particularly interested in the red shaded areas as they represent the branches of the banks that are the largest spatial TFP growth centers. As we would expect, figure 3 reveals that banks which represent the largest spatial TFP growth centers tend to have branches that cluster together in parts of the northeast and in cities elsewhere that are among the largest economies (e.g., Los Angeles and Chicago, to name but a couple). In several smaller major state cities, however, a good example being Indianapolis, there are at best only modest clusters of branches pertaining to banks that act as large spatial TFP growth centers. Having established from figure 1 that the key driver of ΔTFP^{Ind} is the ΔRTS^{Ind} component, in figure 4 we

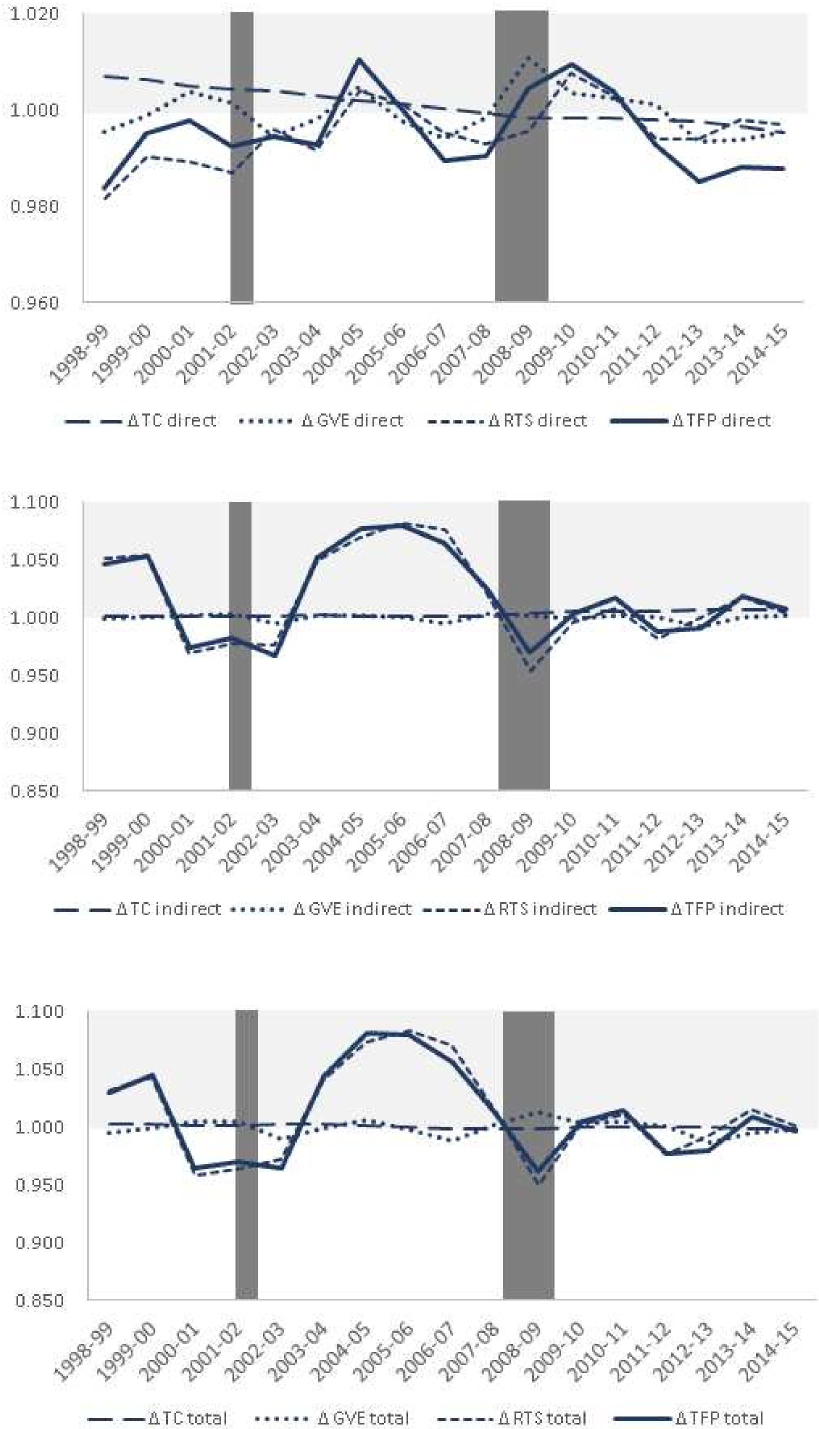

Figure 1: Annual direct, indirect and total TFP growth decompositions for the sample

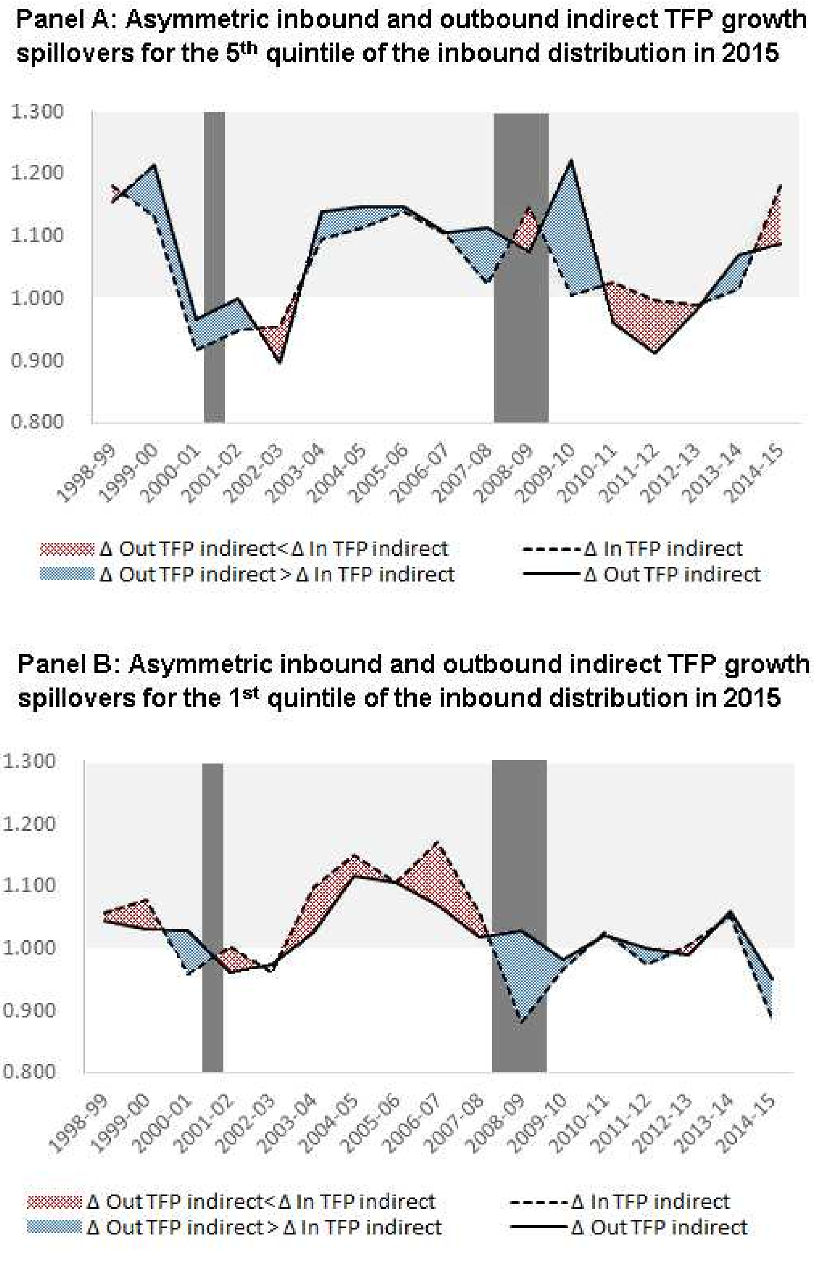

Figure 2: Annual asymmetric bidirectional indirect TFP growth spillovers for selected quintiles

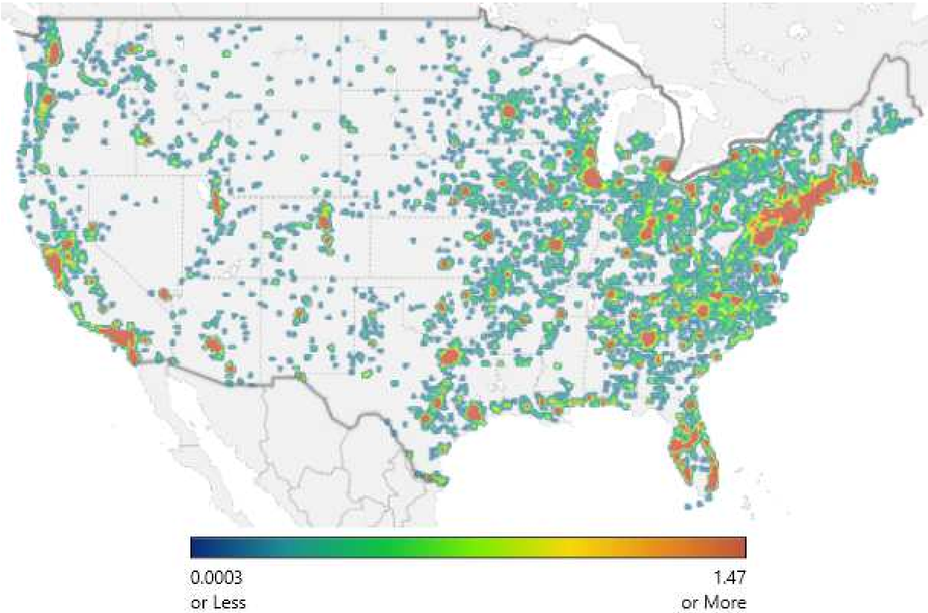

Figure 3: Net generators of TFP growth spillovers in 2015

present for 2015 the geographical distribution of the branches of the banks that are net generators of RTS growth spillovers. As figures 3 and 4 are clearly very similar, we conclude for 2015 that being a net generator of RTS growth spillovers is what determines whether a bank is a spatial TFP growth center.

The spatial TFP growth centers in figure 3 represents one part of the analysis of the net TFP growth spillovers. This is because if there are banks that are spatial TFP growth centers there must be another group of banks that derive net benefits from these centers. In figure 5 we focus on the other part of the analysis as we present the geographical distribution of the branches of the banks that are net recipients of TFP growth spillovers in 2015, i.e., $\Delta TFP^{Ind}_{Out,j} - \Delta TFP^{Ind}_{In,i} < 0$. In this figure we are particularly interested in the dark shaded areas as they represent the branches of the banks that are the largest net recipients of TFP growth spillovers. Comparing figures 3 and 5, we can see, as we would expect, that the large net recipients tend to have branches that cluster near the branches of the banks that act as large spatial TFP growth centers. This is less true in the northeast as the large net recipient banks do not tend to have branches that cluster in this region. The large center banks with branches that cluster in the northeast may instead be large spatial TFP growth centers because they have branches that are near an international financial center.

As the magnitudes of the net TFP growth spillovers in the red and dark shaded areas in figures 3 and 5 are substantive, we examine these net spillovers in more detail by providing an insight into the banks that are the largest net generators and largest net recipients. Recall that Berger and Roman (2017) classify a U.S. bank with total assets between \$1 billion and \$3 billion in 2015 as medium-sized, and a bank with total assets greater than \$3 billion in 2015 as large. Based on this bank size classification, in 2015 the largest 15 net generators and the largest 15 net recipients are an interesting mix of large and medium-sized banks. In summary, both bank size categories are well represented in the largest 15 net generators as there are 7 medium-sized banks in this group. There are fewer medium-sized banks in the largest 15 net recipients as this group includes 9 large banks. Among these groups is at least one bank with total assets in the top 5 in the industry in 2015 as JP Morgan Chase is one of the 15 largest net recipients, and Bank of America and Wells Fargo are among the 15 largest net generators. The smallest banks in the largest 15 net recipients are Woori America Bank (\$1.45 billion) and First Bank and Trust Company (\$1.50 billion), where their total assets in 2015 are in parentheses. In the

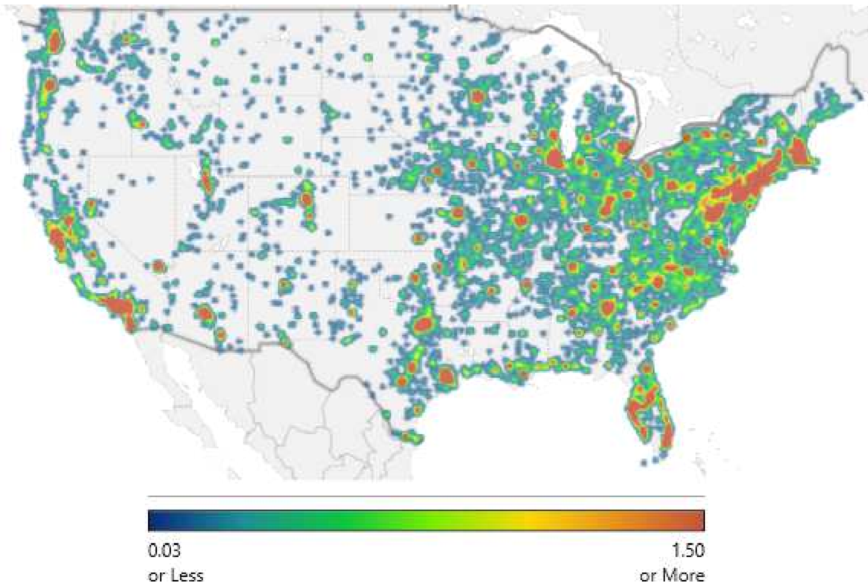

Figure 4: Net generators of returns to scale growth spillovers in 2015

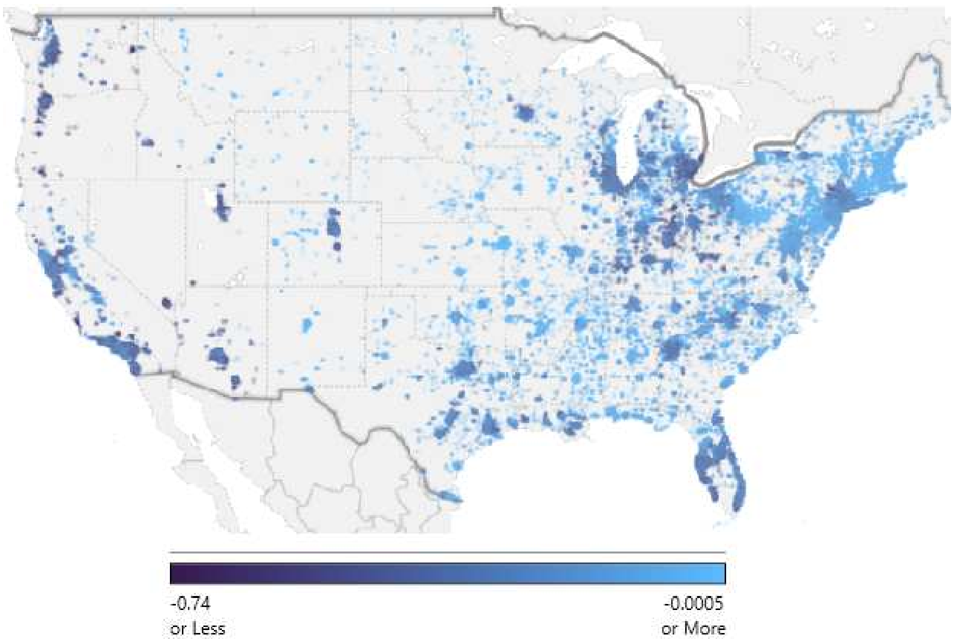

Figure 5: Net recipients of TFP growth spillovers in 2015

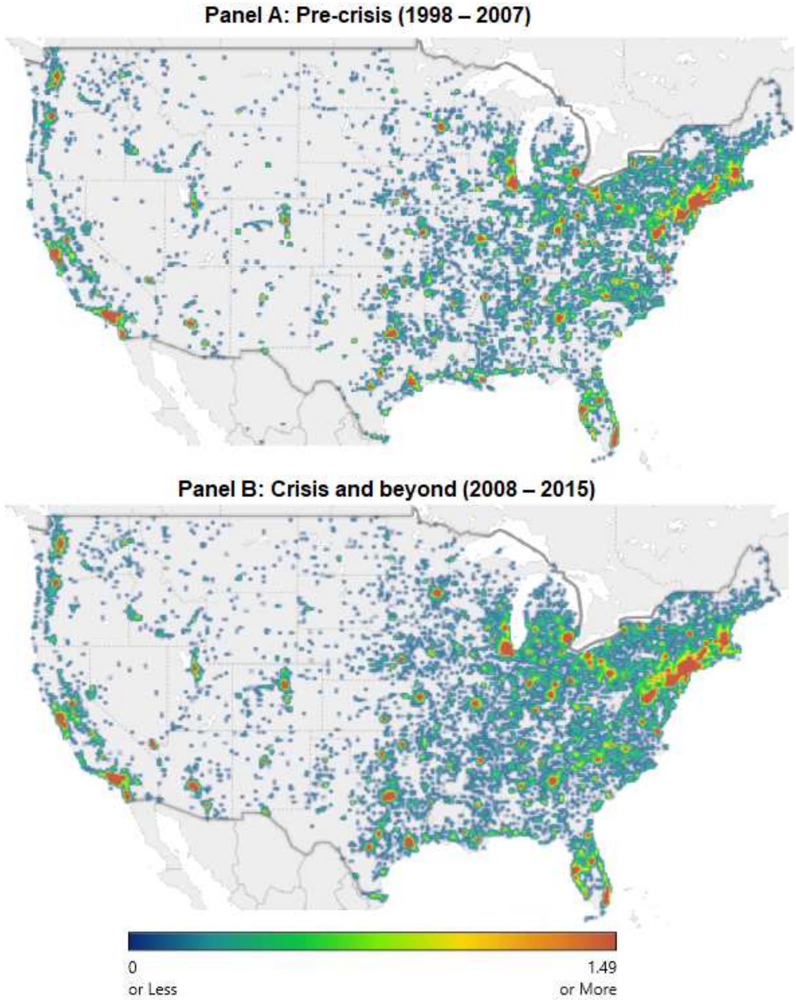

Figure 6: Net generators of TFP growth spillovers over the two subperiods

largest 15 net generators the smallest banks are- Midwest Bank ($1.64 billion); Bear State Bank ($1.46 billion); Peoples Bank ($1.48 billion); First National Bank Texas ($1.41 billion); and Moody National Bank ($1.09 billion). Such findings are informative for U.S. banking authorities as they provide new insights into the different sizes of banks that play key roles in the transmission of productivity spillovers.

Using figure 6 we examine to what extent the large spatial TFP growth centers in 2015 (figure 3) are similar to those for other portions of our study period. In particular, in panels A and B of figure 6 we present the geographical distributions of the branches of the banks that are, on average, classified as spatial TFP growth centers over two subperiods- pre-crisis (1998 – 2007) and crisis and beyond (2008 – 2015). To construct figure 6 we use the branch networks for what is approximately the median year in each subperiod (2002 and 2011).

Comparing the locations of the branches in figures 3 and 6 of the banks that act as large spatial TFP growth centers, where these locations are represented by red shaded areas, gives rise to two key findings. The first is that throughout our study period the banks that act as large spatial TFP growth centers have branches that tend to cluster together in the same locations (the northeast and cities elsewhere with large economies). This indicates that figure 3 is a robust reflection of the branch locations of the banks that act as large spatial TFP growth centers in other portions of our study period. Our second key finding is the increase in the size of the red shaded areas in figure 3, vis-à-vis those in figure 6 (particularly those in panel A). In other words, over our study period we observe an increase in the clustering of the branches of the banks that act as large spatial TFP growth centers, which indicates that these clusters are becoming an increasingly prominent feature of the industry. As we observed from figure 4 for 2015, for both subperiods we find that being a large net generator of RTS growth spillovers is what determines whether a bank represents a large spatial TFP growth center.[21]

5 Concluding Remarks

Our quantitative analysis of productivity centers is the first of its type. The two closest related literatures to our analysis are the quantitative studies of hubs and a study by Pesaran and Yang (2019). The quantitative literature on hubs has thus far focused on uncovering the optimal locations of hubs in transportation and logistics (T&L) networks, or on estimating the impact of hubs on a particular variable when the hubs are selected in advance of the modeling because they are well-known. We add a new dimension to this literature because our paper does not fit into either of these two strands, as we do not consider T&L networks or analyze the effect of productivity centers that are easily observed and can therefore be selected in advance of the modeling. Instead we consider a large number of firms in an industry when it is not immediately obvious which firms act as spatial economic performance centers, which is often the situation in practice. Our principal contribution therefore is to set out the first method to sift through a large number of firms to uncover which firms act spatial TFP growth centers. We use spatial TFP growth as our spatial measure of economic performance as TFP is widely recognized as a comprehensive measure that incorporates both demand and supply changes. A feature of our method is that it is not confined to spatial TFP growth and can be easily applied to any spatial measure to uncover firms that represent various types of spatial center. For example, one could estimate a spatial model of the number of patents for a sample of hi-tech firms and use the patent spillovers from the model to uncover which firms represent technical innovation centers.

There are parallels between our analysis and Pesaran and Yang's study, but whereas our approach operates at the micro level of a large number of firms, their method focuses on the macro level. In particular, they analyze how the sectors that make up a country's production network affect its aggregate output. Their model is based on each sector using the outputs of other sectors as intermediate inputs, but in the context of our micro setting this may not be appropriate for many industries. This is because firms may not use the outputs of the other firms' in the industry as inputs, and in industries where they do, data on this interfirm economic interaction may not be in the public domain. For example, banks interact in this way through interbank lending, but data on such lending is not publicly available, which necessitates that we use our alternative method.

By developing and applying our method to uncover spatial TFP growth centers we make three further contributions. The first two are methodological and are extensions of the method in G&K for their spatial TFP growth decomposition, while the third relates to our application. First, as a spatial stochastic frontier model must first be estimated before one can implement a spatial TFP growth decomposition, rather than follow G&K and estimate a spatial stochastic frontier model with a single SAR variable, we introduce a more sophisticated model that contains multiple SAR variables. Via the multiple SAR variables our new model accounts for spatial interaction in multiple networks, which is likely to be more representative of what is actually the case because a firm may often simultaneously belong to a number of networks. Second, we obtain asymmetric bidirectional spatial TFP growth decompositions. We do so by complementing the unidirectional decomposition in G&K, where the spillover components are spill-ins to a firm, with the corresponding decomposition that includes spill-out components. Third, we

[21] For brevity we do not present the corresponding maps of the branch locations of the banks that are net generators of RTS growth spillovers over the two subperiods. These maps are available from the corresponding author on request.

complement G&K's application of their unidirectional decomposition to U.S. banks using a spatial cost frontier, by focusing on the income side of the industry via bidirectional decompositions from a spatial revenue frontier.

From our application we find that the banks which act as large spatial TFP growth centers tend to have branches that cluster together in the same areas throughout our study period. As we would expect, these areas are the northeast and cities elsewhere that are among the largest economies (e.g., Los Angeles and Chicago). Over our study period we find that the clusters of branches of the banks that act as large spatial TFP growth centers increase in size, which indicates that these clusters are becoming an increasingly prominent feature of the industry. We do, however, observe several smaller major state cities where at best the clusters of branches of the banks that act as large spatial TFP growth centers are modest. We cited Indianapolis as a good example where the large center banks have at best a modest cluster of branches, which is also what we observe in Pittsburgh, Tulsa, El Paso, Columbus (OH), Cleveland (OH) and Albuquerque.

References

Aigner, D., C. A. K. Lovell and P. Schmidt (1977): 'Formulation and estimation of stochastic frontier production function models'. *Journal of Econometrics*, vol. 6, pp. 21-37.

Alumur, S. and B. Y. Kara (2008): 'Network center location problems: The state of the art'. *European Journal of Operational Research*, vol. 190, pp. 1-21.

Anselin, L. (1988): *Spatial Econometrics: Methods and Models*. Dordrecht: Kluwer.

Anselin, L. (2003): 'Spatial externalities, spatial multipliers and spatial econometrics'. *International Regional Science Review*, vol. 26, pp. 153-166.

Anselin, L., A. Varga and Z. Acs (1997): 'Local geographic spillovers between university research and high technology innovations'. *Journal of Urban Economics*, vol. 42, pp. 422-448.

Autant-Bernard, C. and J. P. LeSage (2011): 'Quantifying knowledge spillovers using spatial econometric models'. *Journal of Regional Science*, vol. 51, pp. 471-496.

Battese, G. E. and T. J. Coelli (1988): 'Prediction of firm-level technical efficiencies with a generalized frontier production function and panel data'. *Journal of Econometrics*, vol. 38, pp. 387-399.

Berger, A. N. and C. H. S. Bouwman (2013): 'How does capital affect bank performance during financial crises?' *Journal of Financial Economics*, vol. 109, pp. 146-176.

Berger, A. N. and R. A. Roman (2017): 'Did saving Wall Street really save Main Street? The real effects of TARP on local economic conditions'. *Journal of Financial and Quantitative Analysis*, vol. 52, pp. 1827-1867.

Bhat, C. R. (2001): 'Quasi-random maximum simulated likelihood estimation of the mixed multinomial logit model'. *Transportation Research Part B: Methodological*, vol. 35, pp. 677-693.

Borenstein, S. (1989): 'Centers and high fares: Dominance and market power in the U.S. airline industry'. *Rand Journal of Economics*, vol. 20, pp. 344-365.

Brueckner, J. K., N. J. Dyer and P. T. Spiller (1992): 'Fare determination in airline center-and-spoke networks'. *Rand Journal of Economics,* vol. 23, pp. 309-333.

Campbell, J. F. and M. E. O'Kelly (2012): 'Twenty-five years of center location research'. *Transportation Science*, vol. 46, pp. 153-169.

Chandra, A. and D. Staiger (2007): 'Productivity spillovers in health care: Evidence from the treatment of heart attacks'. *Journal of Political Economy,* vol. 115, pp. 103-140.

Christensen, L. R., D. W. Jorgenson and L. J. Lau (1973): 'Transcendental logarithmic production frontiers'. *Review of Economics and Statistics*, vol. 55, pp. 28-45.

Coelli, T. J., S. Perelman and E. Romano (1999): 'Accounting for environmental influences in stochastic frontier models: With application to international airlines'. *Journal of Productivity Analysis*, vol. 11, pp. 251-273.

Cohen W. M., R. R. Nelson and J. P. Walsh (2002): 'Links and impacts: The influence of public research on industrial R&D'. *Management Science*, vol. 48, pp. 1-23.

Contreras, I., J-F. Cordeau and G. Laporte (2011): 'Stochastic uncapacitated center location'. *European Journal of Operational Research*, vol. 212, pp. 518-528.

Corrado, L. and B. Fingleton (2012): 'Where is the economics in spatial econometrics?' *Journal of Regional Science*, vol. 52, pp. 210-239.

Debarsy, N. (2012): 'The Mundlak approach in the spatial Durbin panel data model'. *Spatial Economic Analysis*, vol. 7, pp. 109-131.

Dick, A. A. (2006): 'Nationwide branching and its impact on market structure, quality, and bank performance'. *Journal of Business*, vol. 79, pp. 567-592.

Diewert, W. E. and K. J. Fox (2017): 'Decomposing productivity indexes into explanatory factors'. *European Journal of Operational Research*, vol. 256, pp. 275-291.

Druska, V. and W. C. Horrace (2004): 'Generalized moments estimation for spatial panel data: Indonesian rice farming'. *American Journal of Agricultural Economics*, vol. 86, pp. 185-198.

Elhorst, J. P. (2009): *Spatial panel data models.* In the *Handbook of Applied Spatial Analysis,* Fischer, M. M., and A. Getis (Eds). New York: Springer.

Filippini, M. and W. Greene (2016): 'Persistent and transient productive inefficiency: A maximum simulated likelihood approach'. *Journal of Productivity Analysis,* vol. 45, pp. 187-196.

Fisher, R. W. and H. Rosenblum (2012): 'Vanquishing too Big to Fail'. *2012 Annual Report of the Federal Reserve Bank of Dallas.* Dallas, TX: Federal Reserve Bank, pp. 5-10.

Fromhold-Eisebith, M. and C. Werker (2013): 'Universities' functions in knowledge transfer: A geographical perspective'. *Annals of Regional Science*, vol. 51, pp. 621-643.

Girma, S. and K. Wakelin (2007): 'Local productivity spillovers from foreign direct investment in the U.K. electronics industry'. *Regional Science and Urban Economics*, vol. 37, pp. 399-412.

Glass, A. J. and K. Kenjegalieva (2019): 'A spatial productivity index in the presence of efficiency spillovers: Evidence for U.S. banks, 1992-2015'. *European Journal of Operational Research*, vol. 273, pp. 1165-1179.

Glass, A. J., K. Kenjegalieva and J. Paez-Farrell (2013): 'Productivity growth decomposition using a spatial autoregressive frontier model'. *Economics Letters*, vol. 119, pp. 291-295.

Glass, A. J., K. Kenjegalieva and R. C. Sickles (2016): 'A spatial autoregressive stochastic frontier model for panel data with asymmetric efficiency spillovers'. *Journal of Econometrics*, vol. 190, pp. 289-300.

Glass, A. J., K. Kenjegalieva, R. C. Sickles and T. G. Weyman-Jones (2018a): 'The spatial efficiency multiplier and common correlated effects in a spatial autoregressive stochastic frontier model'. RISE Working Paper No. 18–003, Houston, TX: Rice Initiative for the Study of Economics, Rice University.

Glass, A. J., A. Kenjegaliev and K. Kenjegalieva (2018b): 'Spatial scale and product mix economies in U.S. banking with simultaneous spillover regimes'. Mimeo.

Glass, A. J., K. Kenjegalieva and T. G. Weyman-Jones (2019): 'The effect of monetary policy on bank competition using the Boone index'. Forthcoming in the *European Journal of Operational Research.*

Gouriéroux, C., A. Holly and A. Monfort (1982): 'Likelihood ratio test, Wald test, and Kuhn-Tucker test in linear models with inequality constraints on the regression parameters'. *Econometrica*, vol. 50, pp. 63-80.

Gude, A., I. C. Álvarez and L. Orea (2018): 'Heterogeneous spillovers among Spanish provinces: A generalized spatial stochastic frontier model '. *Journal of Productivity Analysis*, vol. 50, pp. 155-173.

Horrace, W. C., C. F. Parmeter and I. A. Wright (2019): 'Probability statements for stochastic frontier models with spatial errors'. Mimeo.

Ishfaq, R. and C. R. Sox (2011): 'Center location-allocation in intermodal logistic networks'. *European Journal of Operational Research*, vol. 210, pp. 213-230.

Jondrow, J., C. A. K. Lovell, I. S. Materov and P. Schmidt (1982): 'On the estimation of technical inefficiency in the stochastic frontier production function model'. *Journal of Econometrics*, vol. 19, pp. 233-238.

Kao, Y-H. and A. K. Bera (2013): 'Spatial regression: The curious case of negative spatial dependence'. University of Illinois, Urbana-Champaign, Mimeo.

Koetter, M., J. W. Kolari and L. Spoerdijk (2012): 'Enjoying the quiet life under deregulation? Evidence from adjusted Lerner indices for U.S. banks'. *Review of Economics and Statistics*, vol. 94, pp. 462-480.

Kumbhakar, S. C., G. Lien and J. B. Hardaker (2014): 'Technical efficiency in competing panel data models: A study of Norwegian grain farming'. *Journal of Productivity Analysis*, vol. 41, pp. 321-327.

Lee, L-F. and J. Yu (2012): 'Spatial panels: Random components versus fixed effects'. *International Economic Review*, vol. 53, pp. 1369-1412.

LeSage, J. and R. K. Pace (2009): *Introduction to Spatial Econometrics.* Boca Raton, Florida: CRC Press, Taylor and Francis Group.

Mayer, C. and T. Sinai (2003): 'Network effects, congestion externalities, and air traffic delays: Or why not all delays are evil'. *American Economic Review*, vol. 93, pp. 1194-1215.

Orea, L., I. C. Álvarez and T. Jamasb (2018): 'A spatial stochastic frontier model with omitted variables: Electricity distribution in Norway'. *Energy Journal*, vol. 39, pp. 93-116.

Orea, L. (2002): 'Parametric decomposition of a generalized Malmquist productivity index'. *Journal of Productivity Analysis*, vol. 18, pp. 5-22.

Oude Lansink, A., S. Stefanou and T. Serra (2015): 'Primal and dual dynamic Luenberger productivity indicators'. *European Journal of Operational Research*, vol. 241, pp. 555-563.

Pace, R. K. and R. Barry (1997): 'Quick computation of spatial autoregressive estimators'. *Geographical Analysis*, vol. 29, pp. 232-247.

Pesaran, M. H. (2006): 'Estimation and inference in large heterogeneous panels with a multifactor error structure'. *Econometrica*, vol. 74, pp. 967-1012.

Pesaran, M. H. and C. F. Yang (2019): 'Econometric analysis of production networks with dominant units'. Mimeo.

Qu, X. and L-F. Lee (2015): 'Estimating a spatial autoregressive model with an endogenous spatial weight matrix'. *Journal of Econometrics*, vol. 184, pp. 209-232.

Sealey, C. and J. T. Lindley (1977): 'Inputs, outputs and a theory of production and cost at depository financial institutions'. *Journal of Finance*, vol. 32, pp. 1251-1266.

Sun, K., S. C. Kumbhakar and R. Tveterås (2015): 'Productivity and efficiency estimation: A semiparametric stochastic cost frontier approach'. *European Journal of Operational Research*, vol. 245, pp. 194-202.

Takii, S. (2005): 'Productivity spillovers and characteristics of foreign multinational plants in Indonesian manufacturing 1990-1995'. *Journal of Development Economics*, vol. 76, pp. 521-542.

Tsionas, E. G. and S. K. Mallick (2019): 'A Bayesian semiparametric approach to stochastic frontiers and productivity'. *European Journal of Operational Research*, vol. 274, pp. 391-402.

Tsionas, E. G. and P. G. Michaelides (2016): 'A spatial stochastic frontier model with spillovers: Evidence for Italian regions'. *Scottish Journal of Political Economy*, vol. 63, pp. 243-257.

Viton, P. A. (2010): 'Notes on spatial econometric models'. Ohio State Univeristy. Available at: https://www.yumpu.com/en/document/read/3858779/notes-on-spatial-econometric-models-the-ohio-state-university

Wang, W. and L-F. Lee (2013): 'Estimation of spatial panel data models with randomly missing data in the dependent variable'. *Regional Science and Urban Economics*, vol. 43, pp. 521-538.

Appendix: Technical Details of the Model Estimation Procedure

The log-likelihood function for Eq. 4 is as follows.

$$LL = -\frac{NT}{2}\log\left(2\pi\sigma^2_{\varepsilon^\circ_{it}}\right) + T\log\left|\mathbf{I_N} - \delta_1\mathbf{W_1} - ... - \delta_M\mathbf{W_M}\right| - \frac{1}{2\sigma^2_{\varepsilon^\circ_{it}}}\sum_{i=1}^{N}\sum_{t=1}^{T}\left[\begin{array}{c} y_{it} - x'_{it}\beta - \left(\begin{array}{c}\sum_{j=1}^{N} w_{ij1}x'_{jt}\vartheta_1 + ... \\ +\sum_{j=1}^{N} w_{ijM}x'_{jt}\vartheta_M\end{array}\right) - \left(\begin{array}{c}\delta_1\sum_{j=1}^{N} w_{ij1}y_{jt} + ... \\ +\delta_M\sum_{j=1}^{N} w_{ijM}y_{jt}\end{array}\right) - \\ \overline{x}'_i\zeta - \sum_{j=1}^{N} w_{ij1}\overline{x}'_j\xi_1 - ... - \sum_{j=1}^{N} w_{ijM}\overline{x}'_j\xi_M - \varepsilon^\circ_i \end{array}\right]^2, \tag{A1}$$

where $\mathbf{I_N}$ is the $(N \times N)$ identity matrix and $T\log\left|\mathbf{I_N} - \delta_1\mathbf{W_1} - ... - \delta_M\mathbf{W_M}\right|$ represents the contribution to the log-likelihood from the Jacobian of the transformation from ε°_{it} to y_{it}.[22] Following the simpler model with a single SAR variable and as is standard in the spatial literature, the transformation from ε°_{it} to y_{it} accounts for the endogeneity of the M SAR variables (Anselin, 1988, pp. 63; Elhorst, 2009).

Differentiating Eq. A1 with respect to ε°_i, solving this derivative for ε°_i, and substituting in for ε°_i in Eq. A1 yields (after rearranging) the concentrated log-likelihood function for step 1 of the estimation procedure (Eq. A2). This function is with respect to β, ϑ, δ, ζ and ξ, where we use ϑ, δ and ξ to collectively denote $\vartheta_1, ..., \vartheta_M$, $\delta_1, ..., \delta_M$ and $\xi_1, ..., \xi_M$.

$$LL_C\left(\begin{array}{c}\beta,\vartheta,\delta, \\ \zeta,\xi\end{array}\right) = -\frac{NT}{2}\log\left(2\pi\sigma^2_{\varepsilon^\circ_{it}}\right) + T\log\left|\mathbf{I_N} - \delta_1\mathbf{W_1} - ... - \delta_M\mathbf{W_M}\right| - \frac{1}{2\sigma^2_{\varepsilon^\circ_{it}}}\sum_{i=1}^{N}\sum_{t=1}^{T}\left[\begin{array}{c} y^*_{it} - x'^*_{it}\beta - \left(\begin{array}{c}\sum_{j=1}^{N} w_{ij1}x'^*_{jt}\vartheta_1 + ... \\ +\sum_{j=1}^{N} w_{ijM}x'^*_{jt}\vartheta_M\end{array}\right) - \left(\begin{array}{c}\delta_1\sum_{j=1}^{N} w_{ij1}y^*_{jt} + ... \\ +\delta_M\sum_{j=1}^{N} w_{ijM}y^*_{jt}\end{array}\right) - \\ \overline{x}'^*_i\zeta - \sum_{j=1}^{N} w_{ij1}\overline{x}'^*_j\xi_1 - ... - \sum_{j=1}^{N} w_{ijM}\overline{x}'^*_j\xi_M \end{array}\right]^2, \tag{A2}$$

where an asterisk denotes the transformations of y_{it} and x'_{it} into the quasi-differenced variables in Eqs. A3 and A4. The resulting transformed error $\varepsilon^{\circ *}_{it}$ is in square brackets in Eq. A2, where $\overline{x}'^*_i = \frac{1}{T}\sum_{t=1}^{T} x'^*_{it}$ and $\sum_{j=1}^{N} w_{ijm}\overline{x}'^*_j = \frac{1}{T}\sum_{t=1}^{T}\sum_{j=1}^{N} w_{ijm}x'^*_{jt}$.

$$y^*_{it} = y_{it} - (1-\tau)\frac{1}{T}\sum_{t=1}^{T} y_{it}, \tag{A3}$$

$$x'^*_{it} = x'_{it} - (1-\tau)\frac{1}{T}\sum_{t=1}^{T} x'_{it}, \tag{A4}$$

where τ denotes the weight attached to the cross-sectional component of the data and $0 < \tau^2 = \sigma^2_{\varepsilon^\circ_{it}} / \left(T\sigma^2_{\varepsilon^\circ_i} + \sigma^2_{\varepsilon^\circ_{it}}\right) \leq 1$. For the reasons given in subsection 2.1 we rule out $\tau = 0$ to prevent our model collapsing to the fixed effects specification. When $\tau = 0$ we can see that the transformed mean variables get eliminated from Eq. A2 and the $*$ transformation of the other variables becomes the demeaning procedure used to estimate a fixed effects model.

[22] To simplify the notation in Eq. A1 we subsume α° into x'_{it}.

To obtain the estimates of $\delta_1, ..., \delta_M$ we use the concentrated log-likelihood function in Eq. A5.

$$LL_C(\delta_1, ..., \delta_M) = \varpi - \frac{NT}{2} \log\left[(e_0^* - \delta_1 e_1^* - ... - \delta_M e_M^*)' (e_0^* - \delta_1 e_1^* - ... - \delta_M e_M^*)\right] + T \log |\mathbf{I_N} - \delta_1 \mathbf{W_1} - ... - \delta_M \mathbf{W_M}|, \quad (A5)$$

where ϖ is a constant that does not depend on $\delta_1, ..., \delta_M$. e_0^* and $e_1^*, ..., e_M^*$ denote vectors of ordinary least squares (OLS) residuals from regressing y_t^* and $(\mathbf{I_T} \otimes \mathbf{W_1})y_t^*, ...,$ $(\mathbf{I_T} \otimes \mathbf{W_M})y_t^*$ on $\mathbf{Z_t^*} = \left[\mathbf{X_t^*}, \mathbf{W_1X_t^*}, ..., \mathbf{W_MX_t^*}, \overline{\mathbf{X}}^*, \mathbf{W_1}\overline{\mathbf{X}}^*, ..., \mathbf{W_M}\overline{\mathbf{X}}^*\right]$. Here the variables are denoted as stacked cross-sections for $t \in 1, .., T$, $\mathbf{I_T}$ is the $(T \times T)$ identity matrix and $\otimes$ denotes the Kronecker product. As Eq. A5 can only be solved numerically because closed form solutions for $\delta_1, ..., \delta_M$ do not exist, before we begin the iterations of Eq. A5 and in line with the approach of Pace and Barry (1997) for the simpler case of a spatial model with a single SAR variable and no inefficiency component(s), we calculate $\log |\mathbf{I_N} - \delta_1 \mathbf{W_1} - ... - \delta_M \mathbf{W_M}|$ for a large number of combinations of values of $\delta_1, ..., \delta_M \in \left[\frac{1}{\min(f_1^{\min}, ..., f_M^{\min})}, 1\right]$. In particular, we calculate $\log |\mathbf{I_N} - \delta_1 \mathbf{W_1} - ... - \delta_M \mathbf{W_M}|$ for combinations of values of $\delta_1, ..., \delta_M$ based on 0.001 increments over the above feasible range.

Using ρ to collectively denote the β, $\vartheta_1, ..., \vartheta_M$, ζ and $\xi_1, ..., \xi_M$ vectors of parameters and given the estimates of $\delta_1, ..., \delta_M$, the estimators of ρ and $\sigma^2_{\varepsilon^\circ_{it}}$ are as follows.

$$\widehat{\rho} = b_0 - \delta_1 b_1 - ... - \delta_M b_M = \left(\mathbf{Z_t^{*\prime}Z_t^*}\right)^{-1} \mathbf{Z_t^{*\prime}} \begin{bmatrix} y_t^* - \delta_1 (\mathbf{I_T} \otimes \mathbf{W_1}) y_t^* - ... \\ -\delta_M (\mathbf{I_T} \otimes \mathbf{W_M}) y_t^* \end{bmatrix}, \quad (A6)$$

$$\widehat{\sigma}^2_{\varepsilon^\circ_{it}} = \frac{1}{NT} (e_0^* - \delta_1 e_1^* - ... - \delta_M e_M^*)' (e_0^* - \delta_1 e_1^* - ... - \delta_M e_M^*), \quad (A7)$$

where b_0 and $b_1, ..., b_M$ are the slope estimates from the OLS regressions of y_t^* and $(\mathbf{I_T} \otimes \mathbf{W_1})y_t^*, ..., (\mathbf{I_T} \otimes \mathbf{W_M})y_t^*$ on $\mathbf{Z_t^*}$. Given the estimates of ρ and $\sigma^2_{\varepsilon^\circ_{it}}$, we calculate the estimate of τ using the following concentrated log-likelihood function.

$$LL_C(\tau) = -\frac{NT}{2} \log\left[\varepsilon^{\circ *\prime}\varepsilon^{\circ *}\right] + \frac{N}{2} \log \tau^2, \quad (A8)$$

where an element of the vector $\varepsilon^{\circ *}$ is $\varepsilon^{\circ *}_{it}$, which, as we noted above, is given in square brackets in Eq. A2.

We use the following concentrated log-likelihood function to estimate λ_{uv} in step 2, and thus begin the process of splitting $\widehat{\varepsilon}_{it}$ into estimates of the idiosyncratic error $\widehat{v}_{it}$ and own time-varying inefficiency $\widehat{u}_{it}$.

$$LL_C(\lambda_{uv}) = -NT \log \widehat{\sigma}_{uv} + \sum_{i=1}^{N}\sum_{t=1}^{T} \log\left[1 - \Phi\left(\frac{\widehat{\varepsilon}^\circ_{it}\lambda_{uv}}{\widehat{\sigma}_{uv}}\right)\right] - \frac{1}{2\widehat{\sigma}^2_{uv}} \sum_{i=1}^{N}\sum_{t=1}^{T} \widehat{\varepsilon}^{\circ 2}_{it}, \quad (A9)$$

where $\widehat{\varepsilon}^{\circ 2}_{it}$ is from step 1, Φ is the standard normal cumulative distribution function and

$$\widehat{\sigma}_{uv} = \left(\frac{1}{NT} \sum_{i=1}^{N} \widehat{\varepsilon}^{\circ 2}_{it} / \left[1 - \frac{2\lambda^2_{uv}}{\pi (1 + \lambda^2_{uv})}\right]\right)^{\frac{1}{2}}. \quad (A10)$$

Substituting $\widehat{\lambda}_{uv}$ from maximizing $LL_C(\lambda_{uv})$ into Eq. A10 gives the maximum likelihood estimate of σ_{uv}. As noted in subsection 2.2, we then also in step 2 use $\widehat{\lambda}_{uv}$, $\widehat{\sigma}_{uv}$ and the relevant other estimates to calculate the consistent estimate of own time-varying inefficiency u_{it} (conditional on ε_{it}). In particular, we compute $\widehat{u}_{it}$ using the well-established Jondrow *et al.* (1982) estimator.

In step 3 we split $\widehat{\varepsilon}_i$ into own time-invariant inefficiency $\widehat{\eta}_i$ and the random error component of the firm specific effect that is uncorrelated with the regressors $\widehat{\kappa}_i$. We do so by using the corresponding approach to step 2, which involves using $\widehat{\varepsilon}^\circ_i$ from step 1. Accordingly, using the corresponding equation to A9 $LL_C(\lambda_{\eta\kappa})$ is maximized to compute $\widehat{\lambda}_{\eta\kappa}$, which is then used to calculate $\widehat{\sigma}_{\eta\kappa}$ using the corresponding equation to A10. $\widehat{\lambda}_{\eta\kappa}$ and $\widehat{\sigma}_{\eta\kappa}$, among other things, are then used to calculate the consistent estimate of η_i (conditional on ε_i) using the corresponding equation to 5 in subsection 2.2.