CESifo WORKING PAPERS

6993
2018

April 2018

Contagious Exporting and Foreign Ownership: Evidence from Firms in Shanghai Using a Bayesian Spatial Bivariate Probit Model

Badi H. Baltagi, Peter H. Egger, Michaela Kesina

CESifo

Impressum:

CESifo Working Papers
ISSN 2364-1428 (electronic version)
Publisher and distributor: Munich Society for the Promotion of Economic Research - CESifo GmbH
The international platform of Ludwigs-Maximilians University's Center for Economic Studies and the ifo Institute
Poschingerstr. 5, 81679 Munich, Germany
Telephone +49 (0)89 2180-2740, Telefax +49 (0)89 2180-17845, email office@cesifo.de
Editors: Clemens Fuest, Oliver Falck, Jasmin Gröschl
www.cesifo-group.org/wp

An electronic version of the paper may be downloaded
· from the SSRN website: www.SSRN.com
· from the RePEc website: www.RePEc.org
· from the CESifo website: www.CESifo-group.org/wp

CESifo *Working Paper No. 6993*
Category 8: Trade Policy

Contagious Exporting and Foreign Ownership: Evidence from Firms in Shanghai Using a Bayesian Spatial Bivariate Probit Model

Abstract

Whether a firm is able to attract foreign capital and whether it may participate at the export market depends on whether the fixed costs associated with doing so are at least covered by the incremental operating profits. This paper provides evidence that success for some firms in attracting foreign investors and in exporting appears to reduce the associated fixed costs with exporting or foreign ownership in other firms. Using data on 8,959 firms located in Shanghai, we find that contagion and spillovers in exporting and in foreign ownership decisions within an area of 10 miles in the city of Shanghai amplify fixed-cost reductions for both exporting as well as foreign ownership of neighboring firms. Contagion among exporters and among foreign-owned firms, respectively, amplify shocks to the profitability of these activities to a large extent. These findings are established through the estimation of a spatial bivariate probit model.

JEL-Codes: C110, C310, C350, F140, F230, L220, R100.

Keywords: firm-level exports, firm-level foreign ownership, contagion, spatial econometrics, Chinese firms.

Badi H. Baltagi
Department of Economics & Center for Policy Research
Syracuse University
USA – Syracuse, NY 13244-1020
bbaltagi@maxwell.syr.edu

Peter H. Egger
KOF Swiss Economic Institute
ETH Zurich / Switzerland
egger@kof.ethz.ch

Michaela Kesina
KOF Swiss Economic Institute
ETH Zurich / Switzerland
kesina@kof.ethz.ch

March, 2018
The authors gratefully acknowledge helpful comments by the editor in charge (Zhenlin Yang) and two anonymous reviewers as well as by Baomin Dong and Herbert Egger on earlier versions of the manuscript. Moreover, they are indebted to the participants at the 1st Henan Symposium on Development and Institutional Economics, the Annual World Conference of the Spatial Econometrics Association in 2017, the New York Camp Econometrics in 2015, and the European Trade Study Group Meeting in 2015.

1 Introduction

Models of the *new trade theory* suggest that whether firms export or are part of a multinational network depends – apart from market size, product quality, productivity, factor costs, and trade costs – on the incremental fixed costs associated with such activity.[2]

The majority of both theoretical and empirical contributions in this literature assume that exporting and foreign ownership decisions are carried out independently across firms. This is the case in spite of the theoretical proof of relevant local cross-firm network effects towards global market participation in Krautheim (2012) and to broad evidence of interdependence in exports and foreign direct investment at the aggregate (country-pair) level.[3]

This paper illustrates at the micro level that exporting and foreign ownership decisions are affected by the decisions of other firms in a certain geographical neighborhood. In addition, the exporting and foreign ownership decisions are correlated through dependence not only on the same fundamentals but also through the stochastic shocks.

The paper provides empirical evidence of local contagion, network, or spillover effects in determining export and foreign ownership decisions based on firm-level census-type data from the city of Shanghai – one of the most open regions in terms of exporting and foreign ownership in one of the most prosperous exporting nations on the globe, China. We use cross-sectional data for the year 2002, which is particularly interesting

[2]There is broad theoretical support for these generic arguments. For instance, Melitz (2003) and Helpman, Melitz, and Rubinstein (2008) provide models where firms will enter the export market, if the incremental operating profits from exporting exceed the incremental increase in the fixed costs related to it. Markusen (2002), Helpman, Melitz, and Yeaple (2004), and Barba Navaretti and Venables (2006) provide models of vertically- and horizontally-organized multinational firms where foreign ownership emerges, if the incremental profits from a foreign-owned afiliate exceed the incremental increase in the corresponding fixed costs. Earlier research provided evidence that the fixed costs of multinational firm operation (foreign ownership) first-order dominate those of exporting (see Girma, Görg, and Strobl, 2004).

[3]See Behrens, Ertur, and Koch (2012) and Egger and Pfaffermayr (2016) for evidence of interdependence in aggregate bilateral exports. See Baltagi, Egger, and Pfaffermayr (2007) or Blonigen, Davies, Waddell, and Naughton (2007) for evidence of interdependence in aggregate bilateral foreign direct investments.

since many firms started exporting and attracted foreign capital around that year after China's participation in the World Trade Organization in 2001.

For the purpose of identification and estimation, the paper proposes a novel Bayesian model of contagion with multiple binary dependent variables that are determined jointly by specific latent processes which depend stochastically on each other. The resulting empirical model for the question at stake is a spatial bivariate probit model for exporting and foreign ownership. Estimation by Markov Chain Monte Carlo simulation of the model is shown to perform well in finite samples in a simulation study. One advantage of this estimation procedure relative to standard binary choice models is that it can handle processes with cross-sectionally dependent latent variables (such as the latent profitability of exporting and the latent profitability of foreign ownership). Another advantage relative to single-equation spatial binary choice models is that the approach can handle a process with cross-equation interdependence in the stochastic terms.

The empirical model specification is guided by economic theory and permits a quasi-structural interpretation of the estimated parameters. Economic theory suggests that the (latent) profitability of exporting as well as foreign ownership with heterogeneous firms depends on the attainable profit margin, firm-level productivity, factor costs per efficiency unit, trade costs, fixed costs, and demand. The latent profitability is a log-additive function of the aforementioned arguments in a large class of new trade theory models (see Chaney, 2008; Helpman, Melitz, and Rubinstein, 2008).

The results for firm-level data in Shanghai suggest that contagion among exporters and contagion among foreign-owned firms leads to a significant reduction in the fixed costs of exporting as well as of foreign ownership.

Possible channels of these intra-city, contagious fixed-cost-depressing effects are spillovers which may root in an information dissipation across firms, in an explicit learning induced by cross-firm factor flows (of workers and intermediate goods), and in equilibrium effects (e.g., on the prices of goods and factors). Hence, an increase in the profitability of exporting or foreign ownership (e.g., through policy stimuli such as research funding) in conjunction with the contagion of firms is potentially at least as important for the selection of firms into specific types of activity as a proportional direct reduction of the fixed costs associated with the activity.

The remainder of the paper is organized as follows. Section 2 proposes a stylized theoretical model to motivate the empirical analysis. Section 3 outlines the estimation procedure and presents the performance of the estimator with simulated data. Section 4 describes the data and estimation results, and the last section concludes.

2 Theoretical considerations

A large concurrent literature in international economics assumes that revenues, profits, and fixed costs are linearly separable between exporting and domestic sales on the one hand and between affiliates belonging to a multinational firm on the other hand (see, e.g., Chaney, 2008; Helpman, Melitz, and Rubinstein, 2008, for exporting, and Egger and Seidel, 2013, for multinational plants). The reason why some firms export while others do not are productivity-related operating profits that cover or exceed the fixed costs of running an exporting firm. The reason why some firms belong to a multinational firm while others do not are productivity-related operating profits at the level of the affiliate that cover or exceed the fixed costs associated with foreign ownership. Let us denote type-h firms, with $h = \{e, f\}$ where e stands for exporting and f for foreign ownership. Then, y_{hi} denotes a binary indicator variable taking the value 1 if firm i is of type h and zero otherwise. Determinants of the decision to export and the decision to invest in a foreign firm depend on the type-h-specific operating profits for firm i, Φ_{hi} , as well as the type-h-specific fixed costs for firm i, Ξ_{hi}. The profitability of firm i associated with assuming type h is then reflected in the latent variable $Y^*_{hi} = \frac{\Phi_{hi}}{\Xi_{hi}}$, which generates the binary variable y_{hi}

$$y_{hi} = 1\left(\frac{\Phi_{hi}}{\Xi_{hi}} > U_{hi}\right), \tag{1}$$

where $1(.)$ is an indicator function which takes the value 1 if the condition is satisfied, and zero otherwise. U_{hi} is a non-negative random variable with mean one. Using lower-case letters for logs of $\{Y^*_{hi}, \Phi_{hi}, \Xi_{hi}, U_{hi}\}$, we may write the log-transformed latent process as

$$y^*_{hi} = \phi_{hi} - \xi_{hi}, \text{ with } \ y_{hi} = 1(\phi_{hi} - \xi_{hi} > u_{hi}), \tag{2}$$

where u_{hi} is a random variable with mean zero and infinite support.

Generic models of firms of this kind all have in common that neither y^*_{hi} nor u_{hi} is observed, and they often have the following underlying structure. First, log operating profits, ϕ_{hi}, depend upon market power (mark-ups), on efficiency (productivity), on sector-specific or firm-specific factor costs in efficiency units, and on market potential (trade costs and local versus foreign market size). Second, log fixed costs, ξ_{hi}, depend on the factor requirements (e.g., the necessary assets) to set up a firm of type h (see Markusen, 2002). We argue that, apart from sector-specific characteristics determining fixed costs, the profitability of running a firm of type h depends on other firms' profitability from doing so, denoted by $\overline{y}^*_{hi}$. The latter is a weighted average of the profitability of those firms whose leniency towards becoming a type-h firm affects the profitability of firm i. In the literature this type of interdependence is framed as a network effect and established, e.g., in Krautheim (2012).

Collecting the fundamental drivers of a decision to become an h-type firm into the vector z_{hi}, we can write the latent process as

$$y^*_{hi} = z_{hi}\delta_h + u_{hi} = \lambda_h \overline{y}^*_{hi} + x_{hi}\beta_h + u_{hi}, \tag{3}$$

where $z_{hi} = [\overline{y}^*_{hi}, x_{hi}]$, with $\overline{y}^*_{hi}$ being a scalar and x_{hi} being a vector and $\delta_h = [\lambda_h, \beta'_h]'$, with λ_h being a scalar and β_h being a conformable vector. Notice that the coefficient λ_h measures the strength of interdependence and the potential impact of learning of firms in deciding to become a type-h unit. Given a negative shock on fixed costs, $\lambda_h > 0$ would suggest that fixed costs would decline by more than the negative direct shock due to spillovers from other firms about activity h, while the opposite would be true if $\lambda_h < 0$. Ceteris paribus, the larger λ_h, the larger are the spillovers from other firms about activity h.

3 Econometric model

The previous section argued that the fixed costs of adopting strategy h are potentially affected by the latent profitability of other firms with regard to activity h. With such decisions being interdependent across firms, we need to account for contagion or interdependence in the stochastic model. With the interdependence of firms being geographically (or otherwise) bound, we may formulate a spatial stochastic model. With exporting and foreign ownership not being mutually exclusive but interdependent, such a stochastic model should feature interdependence across units as well as interdependence across latent processes per unit which generate the binary decisions of becoming an exporter or not and attracting foreign capital or not. This is accomplished in the subsequent spatial bivariate probit model.

3.1 Model and notation

Let us denote the export decision of firm $i = 1, ..., n$ by y_{ei} and the foreign ownership decision by y_{fi}, again using $h \in \{e, f\}$. We observe these two binary variables as

$$y_{hi} = 1(y^*_{hi} > 0), \qquad y^*_{hi} = \lambda_h \overline{y}^*_{hi} + x_{hi}\beta_h + u_{hi}, \qquad \overline{y}^*_{hi} = \sum_{j=1}^{n} w_{ij} y^*_{hj}, \tag{4}$$

where $\lambda_h \overline{y}^*_{hi}$ for $h \in \{e, f\}$ summarizes the effect of i's neighboring h-type firms on i's latent profitability associated with strategy h. w_{ij} is a normalized weight describing the strength of the relationship between units i and j, conditional on their location.[4] w_{ij} is positive if two distinct units i and j are neighbors and zero otherwise; it is always zero for $i = j$. Neighborliness can be defined along several lines. In our application we rely on geographical neighborliness between two firms. λ_h denotes the spatial autocorrelation, contagion, interdependence, or spillover parameter for firms of type h, and it will be essential to determine the relative importance of spillovers in the fixed costs associated with activity h.

In general, the vector of covariates x_{hi} is indexed by h, since decisions about exporting and foreign ownership only partly depend on the same exogenous fundamental variables, according to economic theory (see Markusen and Venables, 1998, 2000; Markusen, 2002).

Since the decisions about strategies $h \in \{e, f\}$ depend on an overlapping set of observable fundamentals in x_{hi}, it appears plausible to allow them to be correlated also

[4] It is customary in empirical international economics to take the location of production units as given, while considering their type of activity (here dubbed h) as endogenous (see, e.g., Eaton, Kortum, and Kramarz, 2011; Arkolakis, Ganapati, and Muendler, 2014).

with regard to stochastic shocks. We assume those to be multivariate normal

$$\begin{pmatrix} u_{ei} \\ u_{fi} \end{pmatrix} \sim N\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix}\right), \tag{5}$$

where ρ denotes the tetrachoric correlation between u_{ei} and u_{fi}, and the variances of the errors are normalized to unity (see, for instance, Greene, 2003, for a treatment of the bivariate probit model without accounting for any form of spatial correlation).

Writing the model for all units i and latent outcome h in vector form yields

$$y_h^* = \lambda_h \overline{y}_h^* + x_h \beta_h + u_h, \tag{6}$$

where $\overline{y}_h^* \equiv W y_h^*$, $W = (w_{ij})$ is a row-sum-normalized $n \times n$ matrix reflecting the neighborhood structure.[5] The matrix x_h is $n \times k_h$ where k_h is the number of variables in x_h. u_h is a column vector with n rows.

Stacking the two equations results in

$$\begin{pmatrix} y_e^* \\ y_f^* \end{pmatrix} = \left(\begin{pmatrix} \lambda_e & 0 \\ 0 & \lambda_f \end{pmatrix} \otimes I_n\right) \begin{pmatrix} W & 0 \\ 0 & W \end{pmatrix} \begin{pmatrix} y_e^* \\ y_f^* \end{pmatrix} + \begin{pmatrix} x_e & 0 \\ 0 & x_f \end{pmatrix} \begin{pmatrix} \beta_e \\ \beta_f \end{pmatrix} + \begin{pmatrix} u_e \\ u_f \end{pmatrix}, \tag{7}$$

where I_n denotes an identity matrix of dimension $n \times n$.[6]

3.2 Estimation procedure

Several methods exist to estimate univariate or multivariate binary choice models with conditionally independent data. Among those, maximum likelihood is the most prominent one. The inclusion of the latent profitability associated with strategy h for other firms, $\overline{y}_h^*$, as a determinant of the own latent profitability of strategy h, y_h^*, induces two complications compared to empirical models for independent data: the likelihood function involves an n-dimensional integral, and the reduced form of the latent process is nonlinear. The former relates to computational issues, while the latter relates to the consistency of the model. Ignoring relevant spillovers from $\overline{y}_h^*$ on y_h^* leads to inconsistent

[5] It is customary to normalize W so that a sufficient condition for model stability and an equilibrium is that $|\lambda_h| < 1$.

[6] In principle, W could be specific to activity h.

estimates of the parameters β_h and the corresponding total effects of fundamentals on the probability of adopting strategy h.

For this reason, unless the data-set is manageably small, the standard maximum-likelihood estimator is not feasible with data featuring cross-unit spillovers and interdependence. In fact, for binary choice problems with interdependence by way of inclusion of a weighted average of the dependent latent variable ($\overline{y}_h^*$) on the right-hand side of the empirical model, the following procedures are most commonly used: expectation-maximization methods (see McMillen, 1992); simulated-maximum-likelihood methods through recursive-importance-sampling or integration by simulation using the Geweke-Hajivassiliou-Keane simulator (see Beron and Vijverberg, 2004); generalized-method-of-moments estimation (see Klier and McMillen, 2008); maximum-score estimation (see Lei, 2013); or Markov Chain Monte Carlo (MCMC) simulation (see LeSage, 2000, LeSage and Pace, 2009). These methods have hitherto mostly been applied to univariate (single-equation) problems.

In this paper, we formulate a bivariate binary choice model with spatial dependence, using MCMC techniques to analyze the empirical choice of exporting and foreign ownership for firms. With multiple spatially dependent latent variables and correlation of the stochastic terms across equations, the standard maximum likelihood estimator would be even more difficult to apply. Clearly, any one of the aforementioned methods being able to tackle single-equation binary choice problems with interdependent data would be suitable to analyze multivariate binary choice problems as well. In any case, we are not aware of work that extends the above methods to the multivariate binary choice model with spatial dependence, which is at the heart of this paper's interest.

Latent variable treatment and generic posterior distribution

According to Albert and Chib (1993), who investigated a nonspatial probit model, $p(\beta|y^*) = p(\beta|y^*, y)$. The distribution of the parameters conditional on both y^* and y is the same as the one that only conditions on y^*. In a Bayesian framework, working with latent variables has two nice features: one can sample them, and conditioning on them yields simpler distributions.

In the model of interest here, we may subsume all parameters of interest in θ which contains ρ as well as the elements of y_h^*, β_h, and λ_h for all $h \in \{e, f\}$. Moreover, we may refer to all the data by D, which contains all unique elements of $X = \begin{pmatrix} x_e & 0 \\ 0 & x_f \end{pmatrix}$, W, and y_h for all $h \in \{e, f\}$. Applying Bayes' rule and reformulating yields the following joint posterior distribution

$$\begin{aligned} p(\theta|D) &\propto p(D|\theta) \cdot p(\theta) \\ &\propto p(y_e, y_f|\theta, X, W) p(y_e^*, y_f^*|\beta_e, \beta_f, \lambda_e, \lambda_f, \rho, X, W) \\ &\quad\; p(\beta_e)p(\beta_f)p(\lambda_e)p(\lambda_f)p(\rho), \end{aligned}$$

where we assume independence of the priors in the last line. Since the joint posterior distribution turns out to be intractable, we follow Gelfand and Smith (1990) and calculate the conditional distribution of each parameter θ_ℓ conditional on all the other parameters, $\theta_{-\ell}$ and the data, D. Since we condition on the data throughout, we suppress D among the conditioning arguments in most of what follows for notational simplicity.

Priors

Under independence of the priors of all parameters, the prior distributions may be assumed to be

$$\begin{aligned} \beta_h &\sim N(\underline{\beta}_h, \underline{V}_h) \quad \text{where} \quad \underline{\beta}_h = 0_{k\times 1} \quad \text{and} \quad \underline{V}_h = I_k \cdot 1e^{12}, \\ \lambda_h &\sim U(-1, 1), \\ \rho &\sim U(-1, 1). \end{aligned}$$

These priors are relatively uninformative, reflecting a large degree of uncertainty about the parameters. Intuitively, in calculating the posterior distribution less weight is placed on the prior and more on the data as a consequence.

Likelihood

It will turn out useful to define $L_g = I_n - \lambda_g W$ as well as $\widetilde{L}_g = L_g^{-1}$. The likelihood is stated in terms of the latent variables y_h^* for $h \in \{e, f\}$, and the joint distribution of

(y_e^*, y_f^*) is given by

$$\begin{pmatrix} y_e^* \\ y_f^* \end{pmatrix} \sim N\left(\begin{pmatrix} \widetilde{L}_e x_e \beta_e \\ \widetilde{L}_f x_f \beta_f \end{pmatrix}, \begin{pmatrix} \widetilde{L}_e & 0 \\ 0 & \widetilde{L}_f \end{pmatrix} \left[\begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix} \otimes I_n \right] \begin{pmatrix} \widetilde{L}_e & 0 \\ 0 & \widetilde{L}_f \end{pmatrix}' \right).$$

This yields the likelihood

$$p(y_e^*, y_f^* | \beta_e, \beta_f, \lambda_e, \lambda_f, \rho, X, W) = \frac{1}{2\pi^n} |L_e||L_f||\Sigma|^{-n/2} exp\left[-\frac{1}{2} trace\left(R\Sigma^{-1} \right) \right],$$

where, under the present assumptions, $\Sigma = \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix}$ and $R = \begin{pmatrix} r_{ee} & r_{ef} \\ r_{fe} & r_{ff} \end{pmatrix}$ is a 2×2 matrix containing the elements $r_{gh} = (L_g y_g^* - x_g \beta_g)'(L_h y_h^* - x_h \beta_h)$ with $g, h \in \{e, f\}$.

Conditional distribution of y_e^* and y_f^*

The posterior distributions for the latent variables are calculated using the joint distribution of (y_e^*, y_f^*). Using $g, h \in \{e, f\}$ for $g \neq h$, the conditional distribution of y_g^* given the other parameters is given by[7]

$$y_g^* | \theta_{-y_g^*} \sim N\left(\widetilde{L}_g x_g \beta_g + \rho \widetilde{L}_g L_h \left(y_h^* - \widetilde{L}_h x_h \beta_h \right), (1 - \rho^2) \widetilde{L}_g \widetilde{L}_g' \right).$$

When taking draws of y_e^* and y_f^*, we sample them by applying the method of Geweke (1991),[8] accounting for the state of the observed binary variables y_e and y_f, respectively: we take draws from a right-truncated normal (if the observed binary variable is 0) or from a left-truncated normal (if the observed binary variable is 1). As outlined in LeSage and Pace (2009), every element of the vector y_g^* is drawn separately, taking the interdependence of the units i in the system into account.[9]

Conditional distribution of $\boldsymbol{\beta}_e$ and $\boldsymbol{\beta}_f$

Using $\{g, h\} \in \{e, f\}$ for $g \neq h$, the conditional distribution of β_g given the other parameters is given by

$$\beta_g | \theta_{-\beta_g} \quad \propto \quad N(\overline{\beta}_g, \overline{V}_g),$$

[7]Consistent with the earlier notation, we refer to the distribution of y_g^* conditional on all parameters except y_g^* as $y_g^* | \theta_{-y_g^*}$, here.

[8]In a spatial context, the procedure is outlined in LeSage and Pace (2009).

[9]I.e., when drawing y_{gi}^*, we condition on all parameters other than y_{gi}^* in the system – namely also on all elements of y_g^* except for y_{gi}^*.

where

$$\overline{\beta}_g = \left(\underline{V}_g^{-1} + 1/(1-\rho^2)(x_g' x_g)\right)^{-1} \left[\underline{V}_g^{-1}\underline{\beta}_g + 1/(1-\rho^2)x_g' \left(L_g y_g^* - \rho(L_h y_h^* - x_h \beta_h)\right)\right]$$
$$\overline{V}_g = \left(\underline{V}_g^{-1} + 1/(1-\rho^2)(x_g' x_g)\right)^{-1}.$$

and $\underline{\beta}_g$ and $\underline{V}_g$ are the values of the corresponding prior. Gibbs sampling is applied for drawing β_g.

Conditional distribution of $\boldsymbol{\lambda}_e$ and $\boldsymbol{\lambda}_f$

Using $h \in \{e, f\}$, the conditional distribution of λ_h is given by

$$\lambda_h|\theta_{-\lambda_h} \propto |I_n - \lambda_h W| exp\left[-\frac{1}{2}trace\left(R\Sigma^{-1}\right)\right]. \tag{8}$$

Since this distribution takes an unknown form, we apply a Metropolis-Hastings procedure for simulating it, where we draw a new proposal candidate λ_h' and evaluate the conditional distribution in (8) at both the previous λ_h and the new λ_h' (see LeSage and Pace, 2009).

Conditional distribution of $\boldsymbol{\rho}$

The conditional distribution of ρ is given by

$$\rho|\theta_{-\rho} \propto \begin{vmatrix} 1 & \rho \\ \rho & 1 \end{vmatrix}^{-n/2} exp\left[-\frac{1}{2}trace\left(R\Sigma^{-1}\right)\right], \tag{9}$$

which we also sample by a Metropolis-Hastings approach, since the conditional distribution has an unknown form.

MCMC procedure

With the conditional distributions at hand, we apply a Markov Chain Monte Carlo procedure. After choosing some starting values for the parameters, we draw each parameter from its conditional distribution. Using $h \in \{e, f\}$ and D to denote the data, for each draw, we perform the following steps:

1. Update β_h using Gibbs sampling from its conditional multivariate normal distribution: $\beta_h|\theta_{-\beta_h}, D$.

2. Update ρ using the Metropolis-Hastings procedure.

3. Update λ_h using the Metropolis-Hastings procedure.

4. Update y^*_{hi} using Gibbs sampling from its conditional truncated normal distribution: $y^*_{hi}|\theta_{-y^*_{hi}}, D$.

These steps are repeated until convergence is achieved. Convergence is assessed by means of the Geweke (1992) test and the Raftery and Lewis (1992) I-statistic. The I-statistic should be smaller than 5 and the Geweke statistic should not reject the null hypothesis that the posterior mean of the first 20 percent and the last 50 percent of the draws in the chain are equal.

Monte Carlo simulation study

To illustrate the performance of the bivariate probit model with triangular data, we perform Monte Carlo experiments on a spatial bivariate probit model as it underlies the empirical application in this paper.

In all designs and experiments, we consider two exogenous variables

$$x_{ki} \sim N(0,1) \quad \text{for} \quad k = 2, 3$$

and a constant $x_{1i} = 1$. The three explanatory variables are collected in the vector $x_i = [x_{i1}, x_{i2}, x_{i3}]$ which is the same for the generic latent variables, y^*_1 and y^*_2 (which we use in place of the earlier y^*_e and y^*_f). We assume that the true parameter vectors on the stacked regressors x are $\beta_1 = (1, -2, 1.25)'$ and $\beta_2 = (1, -1, 0.5)'$ for y^*_1 and y^*_2, respectively. Without loss of generality, we assume that all units i are located on a circle for the present design. In particular, the vectors $\overline{y}^*_1$ and $\overline{y}^*_2$ involve a spatial weights matrix which we assume to exhibit a 10-before-10-behind neighborhood structure. Such a neighborhood structure means that unit 1 in equation 1 depends upon an equally-weighted average of y^*_1 of units 2 to 11 as well as of units $n-9$ to n but it will be independent of all other units. The terms $\overline{y}_1^*$ and $\overline{y}_2^*$ will be based on the row-sum-normalized matrix W. With a 10-before-10-behind neighborhood structure, the entries of W are 1/20 for neighbors so that each neighbor to any unit receives a weight of 0.05. The corresponding true parameters on $\overline{y}_1^*$ and $\overline{y}_2^*$, λ_1 and λ_2, respectively, will be specified below. Finally, the stochastic terms (u_1, u_2) are generated as bivariate normal with cross-equation correlation coefficient ρ and unitary variances.

We consider four designs with regard to the configuration of $\{\lambda_1, \lambda_2, \rho\}$: $\{\lambda_1 = 0.4, \lambda_2 = 0.7, \rho = 0.5\}$ (Design 1); $\{\lambda_1 = 0.2, \lambda_2 = 0.3, \rho = 0.5\}$ (Design 2); $\{\lambda_1 = 0, \lambda_2 = 0.7, \rho = 0.5\}$ (Design 3); $\{\lambda_1 = 0.4, \lambda_2 = 0.7, \rho = 0\}$ (Design 4).

For notational convenience, we denote the vector of parameters of interest by $\theta = (\beta_1, \beta_2, \lambda_1, \lambda_2, \rho)$. The latent variables are generated according to the following reduced form:

$$y_h^* \;=\; \widetilde{L}_h x \beta_h + \widetilde{L}_h u_h, \quad h \in \{1, 2\}$$

Using these latent variables, we obtain the observable binary variables $y_h = 1(y_h^* > 0)$. We consider two different sample sizes of $n = \{1,000; 2,000\}$. It will become clear below why even a modest number of eight configurations (two sample sizes and four designs with regard to $\{\lambda_1, \lambda_2, \rho\}$) as considered here is computationally intensive.

Each experiment or configuration is based on 1,000 independent bivariate draws for (u_1, u_2) and, hence, $\overline{y}_1^*$ and $\overline{y}_2^*$ as well as y_1 and y_2. For each replication, we apply the MCMC procedure described in the previous section based on 20,000 draws for the parameter values θ. Hence, with two sample sizes and four designs for $\{\lambda_1, \lambda_2, \rho\}$, 1,000 draws of (u_1, u_2), and 20,000 MCMC draws each, there are 160,000,000 draws. The convergence of the chains is assessed by the Raftery and Lewis (1992) I-statistic and the Geweke (1992) test. In every replication, the first 4,000 MCMC draws were discarded as burn-ins. Due to the presence of high autocorrelation in the draws, we thinned the chain of draws and only kept every 10th draw. After discarding the burn-ins and relying only on the thinned chain of parameters, the remaining 1,600 draws were used to calculate the posterior means of the parameters, $\overline{\theta}$, for every one of the 1,000 replications in each experiment. This yields 1,000 posterior means of the parameter vectors for each experiment. Based on those posterior means, we calculated the summary statistics presented in Tables 1-4 for Designs 1-4, respectively.

– Tables 1-4 about here –

The upper panels of Tables 1-4 contain the results for $n = 1,000$ and the lower panels the results for $n = 2,000$. For all parameters and designs we report the mean, standard deviation, the average bias, and the root mean squared error (RMSE). Furthermore, we report some convergence diagnostics for the individual Markov chains, where we concentrate on the I-statistic proposed by Raftery and Lewis (1992) and the p-values of the Geweke (1992) test. The Monte Carlo analysis is particularly informative with regard to the performance of the estimation procedure given 20,000 draws, 4,000 burn-ins, and a thinning ratio of 1/10. However, we could demonstrate that even using only half of the draws (i.e., 10,000 instead of 20,000) would obtain similar results. This is assuring for the empirical analysis in Section 4, where we will adopt exactly the same approach but with a bigger data-set.

The results in Tables 1-4 suggest that all parameters are estimated well even in modest samples of $n = 1,000$ and $n = 2,000$. The numbers indicate that the small-sample bias is smaller with the larger samples, as expected. But even for $n = 1,000$, the average bias is in the range of less than one percent. The RMSE is only slightly larger than two percent of the true parameter value, but it drops by one-half on average when doubling the sample size. What is most important to us is that the bias and precision of the spillover parameters $\{\lambda_1, \lambda_2\}$ and the cross-equation correlation parameter ρ are as small as those of the other parameters. This makes us confident that the proposed procedure can be fruitfully adopted with multi-equation problems involving cross-sectionally interrelated, binary dependent variables even in moderately-sized samples.

4 Empirical application

4.1 Data set

All data utilized in this study are provided by the National Bureau of Statistics of China (NBS). The data provide information on balance sheets, exports, foreign ownership, and a firm's location for all units with an annual turnover of more than five million Yuan (about 700,000 USD). We focus on manufacturing firms in Shanghai in the year 2002. Shanghai is interesting for the question at stake of the importance of spillovers for becoming an exporter or part of a multinational company because it is one of the biggest cities in China – the biggest exporting country in the world with a lot of foreign direct investment.[10] We consider firms as foreign-owned, if the NBS reports a registration code of 200 or larger for them, so that they are fully- or partly-owned foreign-owned enterprises by mainland-Chinese citizens or other owners abroad or by foreign owners (who may be citizens or entities of Hong Kong, Macao, or Taiwan) or they receive any foreign direct investment from elsewhere than mainland China.

Furthermore, the port of Shanghai is one of the biggest container ports on the globe and has an important role for world market access of China's and, in particular, Shanghai's manufacturers. Indeed, 99% of all goods entering or exiting Shanghai are transferred via the port. Approximately 20% of the foreign trade volume of regions at the Chinese coast are transferred via Shanghai's port.[11]

[10]In this paper, we ignore the potential selection of firms into the considered area itself. Eventually, when not doing so, one might encounter a problem of missing data and of a scale which would go beyond conventional computing capabilities. In some sense, by doing so, our analysis focuses on the short to medium run, where location decisions as such are given, while the nature of firm activity (exporting or not; foreign ownership or not) is in the focus. Moreover, we assume that contagion among firms inside and outside of Shanghai is negligible (otherwise, one would have to consider such contagion even with firms outside of China). Finally, we assume that the focus on the firms with the aforementioned turnover is justified and an associated selection in turnover can be ignored. In the latter regard, we might add that the firms included in the NBS data altogether account for about 95% of China's overall value added in manufacturing between the years 1996 and 2013 (see National Bureau of Statistics of China, 2013).

[11]For more information on the port of Shanghai, see `http://www.shanghaiport.gov.cn/English/introduction/info_001.html`.

Shanghai is a directly-controlled municipality and has the same status as a province. China has a different administrative structure and definition of cities than other countries.[12] In our study, we use manufacturing firms that are located in the main city district of Shanghai (Puxi), firms that are located in the inner suburbs and in the outer suburbs, as well as firms in the rural areas (Shanghai's belt). Table 5 shows the distribution of firms across 26 sectors, and Table 6 illustrates how many firms are exporters versus non-exporters and foreign- versus domestically-owned in the different areas of Shanghai. Table 7 puts the firm numbers for the two states – exporting versus non-exporting and foreign- versus not-foreign-owned in a matrix when pooling all covered sectors and regions of Shanghai.

– Table 5-7 and Figure 1 about here –

Approximately one-third of all firms in the sample are exporters, while about 40% are foreign-owned, with 59% of the foreign-owned units exporting. More than half of all firms are located in inner suburbs, while about one-quarter are located in the outer suburbs. The main city district hosts a lower number of firms than the outer suburbs. The rural area, which consists of the island Chongming, hosts the lowest number of firms among the considered areas of Shanghai. This pattern would be expected, since the center of Shanghai on the one hand benefits from good infrastructure, but on the other hand is characterized by high real estate prices and rents such that setting up plants in this area is expensive. A close proximity to the main city center appears desirable to firms as is obvious from the high number of firms in the inner suburbs surrounding the city center. Among the inner suburbs, Pudong is the biggest host of firms. Specifically, there are more firms in Pudong than in the whole outer suburbs. However, this should not be surprising, since the port of Shanghai is located in this district. The pattern is similar for the two firm types. About 51% of the exporters and and 56% of the foreign-owned firms are located in the inner suburbs. Among those, the probabilities of exporting or being foreign-owned are highest in Pudong.

Figure 1 illustrates the location of exporters and foreign-owned units, respectively,

[12] See Chan (2007) on the definition of cities in China.

in the Shanghai area. By and large, the figure shows a similar pattern of location for the four considered types of firms: foreign-owned exporters, domestically-owned exporters, foreign-owned non-exporters, and domestically-owned non-exporters. There appears to be some general clustering of the firms, but the clustering of exporting firms appears particularly strong within the wider Shanghai area, in particular, when considering foreign-owned exporters.

4.2 Specification

We use the model specification as outlined in Subsection 3.1 for both exporting and foreign ownership. The latent variables for the two activities correspond to their unobserved profitabilities. For the construction of the unnormalized spatial weights matrix, we determine the inverse bilateral haversine distances between all $8,959$ firms in the data.[13] This leads to an $8,959 \times 8,959$ matrix which is symmetric and exhibits zero diagonal entries. We consider spillovers between firms being geographically bound within a radius of 10 miles (we employ alternative specifications of this threshold in the robustness analysis, and the results there suggest that the 10-mile cutoff is selected based on the deviance information criterion). Notice that this is a large radius in view of the fact that (i) the density of firms is high and (ii) bicycles and public transportation are the main means of transport in Shanghai. Moreover, we assume that Chongming is a spatial island where the entries of W between firms in Chongming and other parts of Shanghai are set to zero in spite of their actual distance to each other.

For reasons of interpretation, we row-sum-normalize the weights matrix to obtain the normalized $8,959 \times 8,959$ matrix W and the associated weighted unobserved spillover terms $\overline{y}_e^* = Wy_e^*$ and $\overline{y}_f^* = Wy_f^*$.[14]

[13] The haversine formula is particularly suited for calculating great circle distances between two points i and j on the globe, if two points of location are very close to each other. Denote the haversine function of an argument ℓ by $h(\ell) = 0.5(1 - \cos(\ell))$, and use ϕ_i, ϕ_j, and $\Delta\lambda_{ij}$ to refer to the latitude of i, the latitude of j, and the difference in longitudes between i and j which are all measured in radians. Then, the haversine distance between i and j is defined as $d_{ij} = D \cdot \arcsin(H_{ij}^{1/2})$, where D is the diameter of the globe (e.g., measured in miles) and $H_{ij} = h(\phi_i - \phi_j) + \cos(\phi_i)\cos(\phi_j)h(\Delta\lambda_{ij})$.

[14] Notice that $\overline{y}_e^*$ and $\overline{y}_f^*$ are latent and unobservable. However, we may compute the observable

For a quasi-structural interpretation of the models of y_e^* and y_f^*, the specification of the covariates is key. In line with economic theory, this matrix contains firm-specific and sector-specific covariates besides the constant. At the firm level, we include *Employment*, which is the log of the size of the work force, capturing firm size. *Productivity* is measured as the log of the ratio of total sales to employment. *Intangible asset ratio* is the ratio of intangible assets to total assets and is used as a proxy for knowledge intensity. A larger value of these three variables should raise the profitability and the probability of exporting as well as foreign ownership.[15]

Moreover, we include firm-specific variables to account for geographical factors related to market access and agglomeration. *Distance to port* measures the log haversine distance of a firm to the port of Shanghai, which is important for exporting for several reasons. The most obvious one is Shanghai's geographical location at the coast of China and the importance of the port for global trade. One might in principle consider other modes of transport – such as railroads or motorways. However, those are less important for China's coastal regions, in particular, when it comes to serving customers abroad. As a second geographical covariate, we include *Distance to city center*, which measures the log haversine distance of a firm to the main city center district Huangpu. A closer distance to the center potentially reflects access to better infrastructure, access to specific production factors, access to finance, etc.

Earlier results by Head, Ries, and Swenson (1995), Swenson (2008), Fontagné, Koenig, Mayneris, and Poncet, (2013), and Lovely, Liu, and Ondrich, (2010, 2013) suggest that the agglomeration of firms in regions affects decisions of both exporters and multinational firms. We pay attention to this fact by letting the export and foreign ownership decision

$\overline{y}_e = Wy_e$ and $\overline{y}_f = Wy_f$ for illustrative purposes. In particular, we may consider the correlation coefficients (so-called Moran I statistics) between y_e and $\overline{y}_e$ and y_f and $\overline{y}_f$, respectively. It turns out that those are 0.232 and 0.325, respectively, suggesting somewhat stronger clustering of foreign-owned than exporting units. Notice that this pattern is consistent with the insights gained from Figure 1.

[15] In customary theoretical models in trade, these variables are drawn by firms (productivity and assets) or predetermined by exogenous factors that they stand for in a compact way (employment); see, e.g., Melitz (2003) or Helpman, Melitz, and Rubinstein (2004). Clearly, one could think of these variables as endogenous in different frameworks. However, we refrain from doing so here, as this would further complicate the analysis and raise the demand for data.

in Shanghai depend on the location density of all units in a firm's neighborhood (within 10 miles). The latter is reflected in two covariates. As one measure, the *Average distance to all firms in the neighborhood* (based on log haversine distances) within a radius of 10 miles captures the relative centrality of a firm relative to others, reflecting cheap access to intermediate goods and services from other firms in general (not only exporters or foreign firms). As the other measure, the *Number of other firms in the neighborhood* within a radius of 10 miles captures the relative density of economic activity of firms of any type (domestic sellers versus exporters and domestically-owned versus foreign-owned firms). In general, higher values of distance variables of the aforementioned types reflect bigger distances to the port, the city center, and other firms, respectively, and a bigger number of firms reflects a higher density of economic activity which is not exactly the same as a smaller inverse distance to the average other firm in the neighborhood.

Finally, we include a set of industry-specific covariates. The variable *Sales to profits ratio* (in logs) reflects the profitability or price-cost markup ratio in an industry. We include the variable in a linear and a squared fashion to account for the log-nonlinear impact of the price-cost markups on firm profits in, e.g., monopolistic competition models of heterogeneous firms (see Melitz, 2003).[16] Moreover, we account for the fixed costs associated with exporting and foreign ownership. Specifically, we include the log average total assets of all (other) exporters or foreign-owned firms in the lowest percentile of the respective distribution for each industry. These two variables – *Total assets smallest exporters* and *Total assets smallest foreign firms* – approximate the extent of fixed costs of the marginal firm of either type. The profitability of exporting and/or foreign ownership and, hence, the latent variables y_e^* and y_f^*, depend crucially (and negatively) on these fixed costs, according to economic theory (see Melitz, 2003; Helpman, Melitz, and Yeaple, 2004). All of the considered regressors are exogenous determinants of an individual firm's choice about exporting (see Melitz, 2003; Krautheim, 2012) or foreign ownership (see Helpman, Melitz, and Yeaple, 2004). In order to mitigate a potential bias of these variables associated with the choices of firms, we measure *Total assets smallest*

[16] In fact, the zero-profit conditions for the marginal exporting firm and the marginal foreign-owned firm are nonlinear in the markup in such models.

exporters and *Total assets smallest foreign firms* for each firm i from the distribution of firms which are located outside of a radius of 10 miles of that firm.

– Tables 8 and 9 about here –

Tables 8 and 9 report some descriptive statistics on the covariates for the total sample and for exporting and foreign-owned firms, respectively. First, Table 8 indicates that the average distance of exporters (foreign-owned and domestically-owned) to other firms of their type is smaller than that of non-exporters. Moreover, domestically-owned exporters tend to be situated more closely than foreign-owned ones.

According to Table 9, on average, exporters and foreign-owned firms are bigger in terms of employment than the average firm. This fact is well documented in several studies (see, e.g., Lu, 2010; Ma, Tang, and Zhang, 2012; or Huang, Ju, and Yue, 2013). The productivity of exporters is lower than on average (see Lu, 2010).[17] Foreign-owned firms have a higher productivity than the average firm and also a higher productivity than exporting firms on average. Unlike domestically-owned firms, foreign-owned firms have direct access to technology, knowledge and/or management skills of other units in the same company, which are located abroad. Thus they have an advantage over domestic firms that lack such access. The pattern that foreign-owned firms are more productive than exporting firms is in line with Helpman, Melitz, and Yeaple (2004). Exporters and foreign-owned firms have a higher intangible asset ratio than the average firm.

Somewhat surprisingly, the average distance to the port is slightly higher for exporting and foreign-owned firms than for others. However, it is slightly lower for exporters than for foreign-owned firms. In general, this seems to reflect the trade-off between being closer to the city center and to other firms for factor and technology access versus closer to the port for customer market access. The average sales-to-profit ratio is slightly higher for exporters and for foreign-owned firms compared to other firms. Hence, these firms face tougher competition from their participation at global markets. The smallest ex-

[17] Due to China's strong comparative advantage in labor-intensive sectors compared to developed countries, even less productive firms are competitive on international markets.

porters appear to be somewhat less fixed-cost intensive than the smallest foreign-owned firms in the average manufacturing sector.

4.3 Results

Table 10 presents the main estimation results. In addition to the spatial bivariate probit model, we report nonspatial bivariate probit model results to show the difference in the point estimates (of course, this does not reflect a difference in marginal effects).[18] For both models, we report the parameter estimates and the standard errors in parentheses. For the spatial bivariate probit, we additionally include two convergence diagnostics which we also report with the Monte Carlo simulations, namely Raftery and Lewis' (1992) I-statistic and the p-value of Geweke's (1992) convergence test.[19] For the spatial bivariate probit, we ran 20,000 simulations of which 4,000 were considered as burn-ins. Due to the high presence of autocorrelation, we keep only every 10th iteration. After discarding the respective draws, the MCMC estimates are based on 1,600 draws.

– Table 10 about here –

The results can be summarized as follows. First, the parameter estimates of ρ are 0.611 and 0.309 for the nonspatial and the spatial bivariate probit, respectively. This indicates an interrelation of the exporting and foreign ownership decisions and thus refutes a separate estimation strategy for the two equations by means of an ordinary nonspatial or a single-equation spatial probit model.[20] Disregarding spillovers among firms of the same type results in an upward bias of the cross-equation correlation coefficient.

Second, for the spatial bivariate probit model, Raftery and Lewis' (1992) I-statistic and the p-values of Geweke's (1992) test point to convergence of the Markov chains.

[18]Recall that ignoring a relevant spatially lagged latent dependent variable results in an omitted variables bias.

[19]Recall that the I-statistic should be smaller than 5 and the Geweke test should not reject the null hypothesis.

[20]Ignoring the cross-equation dependence in the disturbances by estimating two separate probit models for exporting and foreign ownership leads to an efficiency loss. We present the results for separate nonspatial and spatial univariate probit models of that kind in Table 17 in Appendix C.

Third, the parameters of the spatially lagged dependent variables are positive and highly significant. Hence, there are cross-firm spillovers in both equations for firms of the same type. Based on these results, the simple, nonspatial bivariate probit model is rejected. In light of the arguments in Section 2, the results point to a statistically significant reduction in the fixed costs of exporting and foreign ownership accruing to spillovers from other firms of the same type. Overall, the economic effect of contagion and spillovers appears to be large.

Fourth, higher *Employment* and higher *Productivity* increase the probability of exporting as well as the probability of foreign ownership. The positive effect of *Productivity* is more pronounced in the foreign ownership equation. To attract foreign capital, firms need to be more productive than for exporting. This is in line with the theoretical arguments of Helpman, Melitz, and Yeaple (2004) and with the evidence on cherry-picking of multinational firms with regard to their foreign investment decisions (see Blonigen, Fontagné, Sly, and Toubal, 2014). A higher *Intangible asset ratio* exhibits a positive impact on both exporting and foreign ownership.

Distance to port does not have a statistically significant impact in either equation of the spatial bivariate probit. In contrast, the nonspatial model does point to a statistically significant impact of this variable. Notice that the reduced form of the latent processes leads to a model which includes spatially weighted explanatory variables. We conjecture that the spatial lag of distance to the port is collinear with distance to the port itself, which explains this difference in the parameter estimates between the spatial and nonspatial bivariate probit models. *Distance to city center* is positive and statistically significant only in the foreign ownership equation. Hence, foreign-owned firms prefer taking ownership of units which are located at some distance of the city center (where real estate prices are relatively high), while this is less the case for exporting firms. A higher *Average distance to other firms in the neighborhood* reduces the probability of exporting and foreign ownership. Proximity to other firms exhibits a positive impact on the probability of exporting and foreign ownership which is in line with the positive estimates of the spatial autocorrelation parameter. Finally, in line with earlier aforementioned work, a greater local density of firms captured by *Number of other firms* *in the neighborhood* raises the export propensity. It also raises the propensity of being foreign owned of the average unit.

The *Sales to profits ratio* in the industry has (hump-shaped) non-linear effects on the linear index underlying the latent profitability of exporting or foreign ownership (the coefficient on the linear term is positive and the one on the squared term is negative).[21] Higher fixed costs, which are measured by *Total assets smallest exporters* and *Total assets smallest foreign firms* in the same industry, exhibit a negative impact on the probability of being an exporter and foreign owned. The decision to export is negatively affected by higher minimum export fixed costs, and the probability of foreign ownership is negatively affected by higher minimum fixed costs of foreign ownership. This is in line with economic theory (see Markusen, 2002; Markusen and Maskus, 2002; Markusen and Venables, 1998, 2000; Helpman, Melitz, and Yeaple, 2004).

A quantitative comparison of the estimates between the spatial and nonspatial models is not possible on the basis of Table 10, but we will relegate it to where we discuss the marginal effects on latent and binary outcomes (see Section 4.5).

4.4 Robustness

In this subsection, we assess the robustness of the findings presented above along two lines. First of all, we check whether state ownership plays a role for the decision to export and for the attractiveness of firms to foreign investors. We do so by adding an indicator variable which is unity whenever a firm is at least partly state-owned.[22] This is important, since state-owned firms in China tend to be less productive than others (see Baltagi, Egger, and Kesina, 2016) which should reduce export market success and their attractiveness to foreign investors. On the other hand, state ownership might relax

[21] Clearly, the direct effect of that variable on the probability of adopting a particular strategy is nonlinear for three reasons: the polynomial form in the linear index; the nonlinear form of the reduced form of the bivariate model accruing to the spillover terms y_e^* and y_f^*; and the nonlinear functional form of the cumulative normal distribution underlying the probit model.

[22] The NBS reports the state-owned capital for each firm. We classify a firm as to be state-owned here, if the share of total assets accruing to state ownership is larger than zero. State ownership and foreign ownership may overlap in a firm.

financial constraints and the need to cover fixed costs of exporting or foreign ownership which would have the opposite effect. The results in Table 11 suggest that, on net, state ownership reduces the propensity to export as well as that of foreign ownership, at least in the sample at hand. However, the other results are qualitatively insensitive to the inclusion of the binary indicator variable *State-owned*.

– Table 11 about here –

Second, we check the robustness of our results by changing the geographical reach of spillovers in the decisions of interest. In the previous subsection, we allowed spillovers to occur within 10 miles. There, all elements of W where the distance was bigger than 10 miles were set to zero. In the robustness checks, we vary this cutoff determining the geographical reach of spillovers by considering alternative values of 12, 8, and 5 miles, respectively.

Table 12 contains the results for the spatial bivariate probit model. The columns involve weights matrices that are based on positive cell entries w_{ij} if firms i and j are closer than 12 miles (W_{12}), 8 miles (W_8), or 5 miles (W_5), respectively.

– Table 12 about here –

The results may be summarized as follows. First, the estimated tetrachoric correlation parameter between the exporting and foreign ownership equation is 0.309 for spillovers within 12, 8, and 5 miles, respectively. These values are virtually identical to the benchmark value reported in Table 10. Second, the spatial autocorrelation parameters vary with the different thresholds. Notice, that W-matrices with a lower threshold have more zero elements than ones based on a bigger threshold. It will generally be the case that the unnormalized weights have the property $\sum_{j=1}^{N} w_{5,ij}^{0} \leq \sum_{j=1}^{N} w_{8,ij}^{0} \leq \sum_{j=1}^{N} w_{12,ij}^{0}$. Therefore, with row-sum normalization the positive individual cells of the respective normalized matrices have the property $w_{12,ij} \leq w_{8,ij} \leq w_{5,ij} \leq 1$. Compared to our results using W_{10}, we find that the spatial autocorrelation parameters are higher (lower) when allowing for a bigger (smaller) geographical reach of spillovers. As in Table 10, we generally find that $\lambda_f > \lambda_e$.

Notice that the deviance information criterion takes on values of $\{43,490; 43,457; 43,446; 43,462\}$ when using the spatial weights matrices $\{W_5; W_8; W_{10}; W_{12}\}$. Hence, among the choices given, we would select the process in Table 10 which is based on a matrix of W_{10}.

4.5 Effect estimates

Clearly, the results in Tables 10-12 are indicative about the qualitative effects of the fundamentals on the response probabilities regarding exporting and foreign ownership. However, the nonlinear nature of the probit model together with the nonlinear structure of the reduced form of the underlying latent processes do not permit a quantitative assessment. Therefore, we devote this subsection to a discussion of estimates of the effects of a change in the k-th covariate on outcome. For this, it will be useful to introduce the one-standard-deviation change in covariate k in equation h, which is scaled by the corresponding parameter estimate, $\widehat{\beta}_{h,k}$, as $\widehat{\Delta}_{h,k}$. The later serves to facilitate a comparison of the regressors in terms of their relative importance for latent outcomes (i.e., the respective profitabilities) and choice probabilities regarding exporting and foreign ownership.

Regarding effect estimates on the latent profitabilities of exporting and foreign ownership let us define for $h \in \{e, f\}$ the vectors

- of <u>d</u>irect effects: $\widehat{d}^d_{h,k} = vecdiag(\widetilde{L}_h)\widehat{\Delta}_{h,k}$
- of effects <u>o</u>n others, $\widehat{d}^o_{h,k} = (\widetilde{L}_h \iota_N)\widehat{\Delta}_{h,k} - \widehat{d}^d_{h,k}$
- of effects <u>f</u>rom others, $\widehat{d}^f_{h,k} = (\widetilde{L}'_h \iota_N)\widehat{\Delta}_{h,k} - \widehat{d}^d_{h,k}$
- of <u>t</u>otal effects, $\widehat{d}^t_{h,k} = \widehat{d}^d_{h,k} + \widehat{d}^f_{h,k}$

where $L_h = (I_N - \widehat{\lambda}_h W)$ and $\widetilde{L}_h = L_h^{-1}$, as above.

While the estimates of $\widehat{d}^\ell_{h,k}$ with $\ell \in \{d, o, f, t\}$ allude to the relative importance of regressors for latent profitabilities, they do not permit immediate conclusions regarding associated changes in choice probabilities due to the nonlinear mapping of the former into the latter. With regard to moments of the effects on choice probabilities for equation $h \in \{e, f\}$, we can use the draws $\widehat{y}^*_h$ and the above estimates $\widehat{d}^\ell_{h,k}$. Then, counterfactual latent outcomes can be calculated as $\widetilde{y}^{*\ell}_{h,k} = \widehat{y}^*_h + \widehat{d}^\ell_{h,k}$ and the corresponding counterfactual binary outcomes are $\widetilde{y}^\ell_{h,k} = (\widetilde{y}^{*\ell}_{h,k} > 0)$. The latter then permits the computation of changes in marginal probabilities as

$$\Delta P^\ell_{h,k} = E(\widetilde{y}^\ell_{h,k}) - E(y_h) \tag{10}$$

and in joint probabilities as

$$\Delta P^\ell_{rs,k} = E(\widetilde{y}^\ell_{e,k} = r, \widetilde{y}^\ell_{f,k} = s) - E(y_e = r, y_f = s) \text{ for } r, s \in \{0, 1\}. \tag{11}$$

Clearly, all of the above is defined and could be done for $\ell \in \{d, o, f, t\}$. However, we focus on estimates of total effects with $\ell = t$ here for the sake of brevity. Table 13 summarizes the corresponding estimates regarding effects on the continuous latent exporting and foreign ownership profitabilities, $\widehat{d}^t_{h,k}$, with the data at hand when using the parameters in Table 10. Since the spatial model entails effects which vary across the units of observation i due to spatial multiplier effects, we report moments of the distribution of the effects for the covariates in that model, while we only report average effects (and their standard errors) for the nonspatial bivariate probit model, where, due to the absence of spillovers, average effects do not vary with the location of firms. For the sake of brevity, let us mostly focus on the average effects which are reported at the bottom of each panel in the table.

– Tables 13-15 about here –

The results in Table 13 suggest the following conclusions. Among the considered regressors, firm size (*Employment*), efficiency (*Productivity*), and the degree of local competition (*Sales to profit ratio*) stand out as drivers of both the gains from exporting and foreign ownership in the spatial model. The fixed costs of exporting (*Total assets of smallest exporters*) and of foreign ownership (*Total assets of smallest foreign firms*) are important obstacles to exporting and foreign ownership, respectively, with the corresponding scaled effects on latent outcomes being among the highest across the regressors in absolute value. The reported moments of effects across the observations i suggest that there is a large degree of heterogeneity in the responses to homogeneous shocks, which accrues to spillovers and the geographical location of firms. In general, the average effects are much smaller in the nonspatial bivariate probit model than in its spatial counterpart, suggesting that the omitted-variables bias associated with an ignorance of the spatially lagged latent variables in the model is of large magnitude.

Table 14 reports on the changes in marginal choice probabilities of exporting and foreign ownership associated with the changes in latent outcomes in Table 13, and Table 15 does the same for the joint choice probabilities. The scaled increase in firm size (*Employment*) raises the propensity of exporting by about 54 percentage points and the one of foreign ownership by more than 41 percentage points for the average firm in the data, when using the spatial model. The scaled change in efficiency (*Productivity*) is also very large with 21 and 50 percentage points on the same marginal choices, respectively, in the spatial model. The scaled increase in exporter fixed costs (*Total assets of smallest exporters*) reduces the propensity of exporting by about 23 percentage points, and the scaled increase in foreign ownership fixed costs (*Total assets of smallest foreign firms*) reduces the propensity of foreign ownership by about 15 percentage points. Most of the estimates on the marginal and joint response probabilities are estimated at relatively high statistical precision, when considering customary thresholds. Across the board, in the nonspatial model the corresponding response-probability changes are much smaller than in the spatial model.

The results in Table 15 suggest that, e.g., an increase in fixed exporting costs reduces the probability of exporting through foreign-owned firms by about 15 percentage points and the one through domestically-owned firms by about 8 percentage points. Moreover, an increase in fixed foreign ownership costs reduces the probability of foreign ownership of both exporting and non-exporting firms by about 8 percentage points. As with the marginal response probabilities, the corresponding effects on the joint response probabilities are much smaller in absolute value with the nonspatial bivariate probit model.

5 Conclusion

This paper focuses on the role of fixed costs for exporting and foreign ownership. The trade literature suggests that whether a firm is able to attract foreign capital and whether it may participate at the export market depends on whether the fixed costs associated with doing so are at least covered by the incremental operating profits. The majority of theoretical and empirical contributions to this literature assume that exporting and foreign ownership decisions are made independently across firms. This paper illustrates that this is not the case, but decisions are interdependent or contagious within a certain geographical neighborhood, and that they are interdependent between exporting and foreign ownership.

For estimation, the paper proposes a Bayesian model of contagion with multiple binary variables that are determined jointly by specific latent processes which depend stochastically on each other. An advantage of this estimation procedure relative to standard binary choice models is that it can handle processes with cross-sectionally dependent latent variables. An advantage relative to single-equation spatial binary choice models is that the approach can handle a process with cross-equation dependence in the stochastic terms and, eventually, even in the dependent variables.

We apply a bivariate probit model to the ability to export and/or attract foreign capital in a sample of 8,959 firms in Shanghai. The results suggest that likely export success for some firms and likely success in attracting foreign investors appears to reduce the associated fixed costs with exporting or foreign ownership in other firms. Contagion in exporting and in foreign ownership within an area of 10 miles in Shanghai leads to an amplification of a reduction in direct fixed costs of firms. A negative direct shock on all firms' fixed costs are inflated due to spillovers and lead to increases in fixed costs which are larger than the direct shocks, which affect firms' globalization strategies.

References

Albert, J.H. and Chib, S. (1993). Bayesian analysis of binary and polychotomous response data. *Journal of the American Statistical Association* **88**(422), 669–679.

Arkolakis, C., Ganapati, S., and Muendler, M. (2014). The extensive margin of exporting products: a firm-level analysis. NBER Working Paper No. 16641.

Baltagi, B.H., Egger, P.H., and Kesina, M. (2016). Firm-level productivity spillovers in China's chemical industry: a spatial Hausman-Taylor approach. *Journal of Applied Econometrics* **31**(1), 214–248.

Baltagi, B.H., Egger, P.H., and Pfaffermayr, M. (2007). Estimating models of complex FDI: are there third-country effects? *Journal of Econometrics* **140**(1), 260–281.

Barba Navaretti, G. and Venables, A.J. (2006). *Multinational Firms in the World Economy.* Princeton University Press.

Behrens, K., Ertur, C., and Koch, W. (2012). Dual gravity: using spatial econometrics to control for multilateral resistance. *Journal of Applied Econometrics* **27**(5), 773–794.

Beron, K.J. and Vijverberg, W.P.M. (2004). Probit in a spatial context: a Monte Carlo analysis. In: Anselin, L., Florax, R.J.G.M., and Rey, S.J. (Eds). Advances in spatial econometrics. Methodology, tools and applications, Berlin, Germany: Springer.

Blonigen, B.A., Davies, R.B., Waddell, G.R., and Naughton, H.T. (2007). FDI in space: spatial autoregressive relationships in foreign direct investment. *European Economic Review* **51**(5), 1303–1325.

Blonigen, B.A., Fontagné, L., Sly, N., and Toubal, F. (2014). Cherries for sales: the incidence and timing of cross-border M&A. *Journal of International Economics*, **94**(2), 341–357.

Chan, K.W. (2007). Misconceptions and complexities in the study of China's cities: Definitions, statistics, and implications. *Eurasian Geography and Economics* **48**(4), 383–412.

Chaney, T. (2008). Distorted gravity: the intensive and extensive margins of international trade. *American Economic Review* **98**(4), 1707–1721.

Eaton, J., Kortum, S., and Kramarz, F. (2011). An anatomy of international trade: evidence from French firms. *Econometrica* **79**(5), 1453–1498.

Egger, P.H., and Pfaffermayr, M. (2016). A generalized spatial error components model for gravity equations. *Empirical Economics* **50**(1), 177–195.

Egger, P.H. and Seidel, T. (2013). Corporate taxes and intra-firm trade. *European Economic Review* **63**, 225–242.

Fontagné, L., Koenig, P., Mayneris, F., and Poncet, S. (2013). Clustering the winners: the French policy of competitiveness clusters. *Journal of Regional Science* **53**(5), 897–922.

Gelfand, A.E., and Smith, A.F.M. (1990). Sampling-based approaches to calculating marginal densities. *Journal of the American Statistical Association* **85**, 398–409.

Geweke, J. (1991). Efficient simulation from the multivariate normal and student t-distributions subject to linear constraints. In: *Computer Sciences and Statistics Proceedings of the 23d Symposium on the Interface*, 571–578.

Geweke, J. (1992). Evaluating the accuracy of sampling-based approaches to calculating posterior moments. In: Bernardo, J.M., Berger, J.O., Dawid, A.P., and Smith, A.F.M. (eds.), *Bayesian Statistics*. Oxford, UK: Clarendon Press.

Geweke, J. (2005). Contemporary Bayesian Econometrics and Statistics. Hoboken, New Jersey : Wiley-VCH.

Greene, W.H. (2003). *Econometric Analysis.* Upper Saddle River, NJ, USA: Prentice-Hall.

Girma, S., Görg, H., and Strobl, E. (2004). Exports, international investment, and plant performance: evidence from a non-parametric test. *Economics Letters* **83**(3), 317–324.

Head, K., Ries, J., and Swenson, D. (1995). Agglomeration benefits and location choice: evidence from Japanese manufacturing investment in the United States. *Journal* *of International Economics* **38**(3–4), 223–247.

Helpman, E., Melitz, M., and Rubinstein, Y. (2008). Estimating trade flows: trading partners and trading volumes. *Quarterly Journal of Economics* **123**, 441–487.

Helpman, E., Melitz, M., and Yeaple, S.R. (2004). Export versus FDI with heterogenous firms. *American Economic Review* **94**(1), 300–316.

Huang, H., Ju, J., and Yue, V.Z. (2013). A unified model of structural adjustments and international trade: theory and evidence from China. Working paper.

Klier, T. and McMillen, D.P. (2008). Clustering of auto supplier plants in the United States: generalized method of moments spatial logit for large samples. *Journal of Business and Economic Statistics* **26**(4), 460–471.

Krautheim, S. (2012). Heterogeneous firms, exporter networks and the effect of distance on international trade. *Journal of International Economics* **87**, 27–35.

Lei, J. (2013). Smoothed spatial maximum score estimation of spatial autoregressive binary choice panel models. Working paper.

LeSage, J.P. and Pace, R.K. (2009). *Introduction to Spatial Econometrics*. Boca Raton, FL, USA: CRC Press/Taylor & Francis.

LeSage, J.P. (2000). Bayesian estimation of limited dependent variable spatial autoregressive models. *Geographical Analysis* **32**(1), 19–35.

Lovely, M.E., Liu, X., and Ondrich, J. (2010). The location decisions of foreign investors in China: untangling the effects of wages using a control function approach. *Review of Economics and Statistics* **92**(1), 160–166.

Lovely, M.E., Liu, X., and Ondrich, J. (2013). Does final market demand elasticity influence the location of export processing? Evidence from multinational decisions in China. *The World Economy* **36**(5), 509–536.

Lu, D. (2010). Exceptional exporter performance? Evidence from Chinese manufacturing firms. Working paper.

Ma, Y., Tang, H., and Y. Zhang (2012). Factor intensity, product switching, and productivity: evidence from Chinese exporters. Working paper.

McMillen, D.P. (1992). Probit with spatial autocorrelation. *Journal of Regional Science* **32**(3), 335–348.

Markusen, J.R. and Venables, A.J. (1998). Multinational firms and the new trade theory. *Journal of International Economics* **46**(2), 183–203.

Markusen, J.R. and Venables, A.J. (2000). The theory of endowment, intra-industry and multinational trade. *Journal of International Economics* **52**(2), 209–234.

Markusen, J.R. (2002). *Multinational Firms and the Theory of International Trade.* Cambridge, Massachusetts, USA: MIT Press.

Markusen, J.R. and Maskus, K.E. (2002). Discriminating among alternative theories of the multinational enterprise. *Review of International Economics* **10**(4), 694–707.

Melitz, M. (2003). The impact of trade on intra-industry reallocations and aggregate industry productivity. *Econometrica* **71**(6), 1695–1725.

National Bureau of Statistics of China (2013). China's annual survey of industrial firms (CASIF): Introduction.

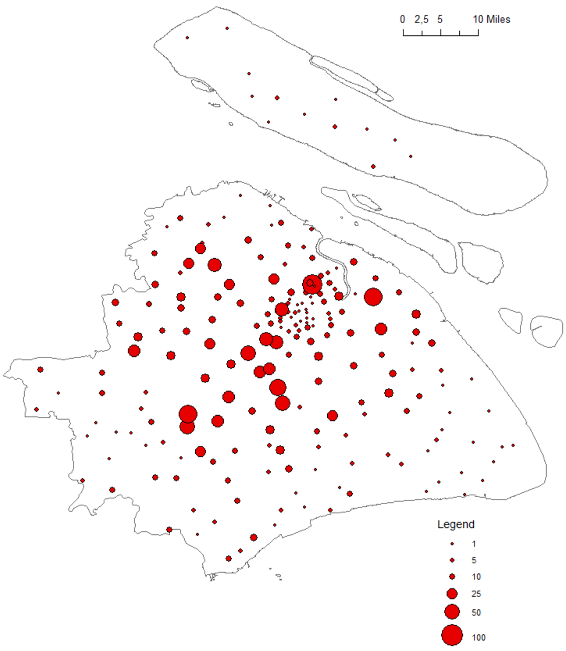
Raftery, A.E. and Lewis, S.M.(1992). How many iterations in the Gibbs sampler? In: Bernardo, J.M., Berger, J.O., Dawid, A.P., and Smith, A.F.M. (eds.), *Bayesian Statistics.* Oxford, UK: Clarendon Press.

Swenson, D. (2008). Multinational firms and new Chinese export transactions. *Canadian Journal of Economics* **41**(2), 596–618.

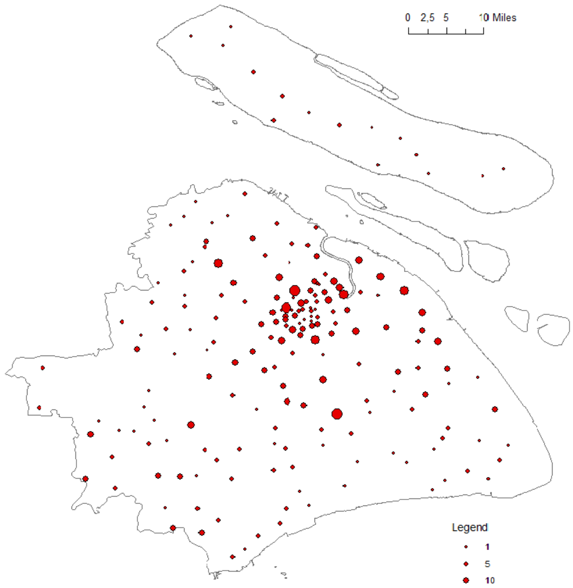
Figures

Figure 1: Location of different firm types in Shanghai

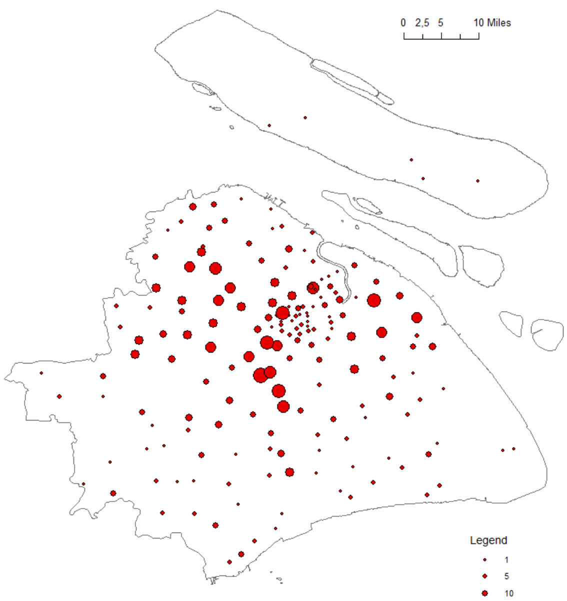
(a) Foreign-owned exporters



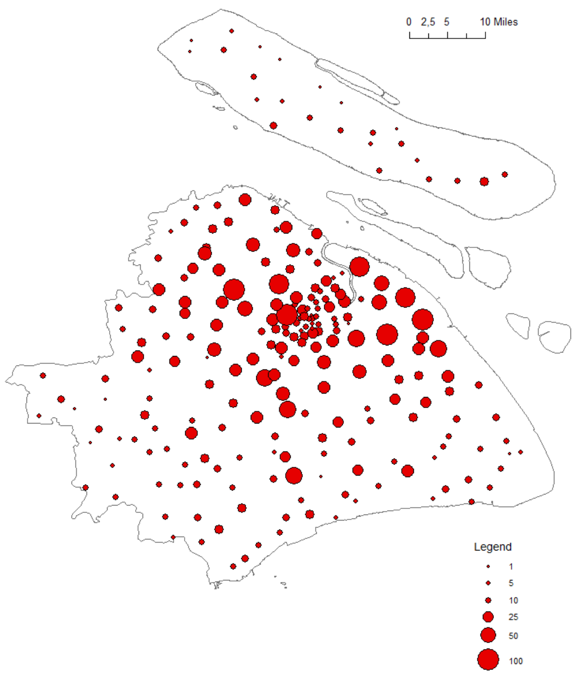
(b) Domestically-owned exporters



(c) Foreign-owned non-exporters



(d) Domestically-owned non-exporters



Tables

Table 1: Monte Carlo Results: Design 1

	β_{11}	β_{12}	β_{13}	β_{21}	β_{22}	β_{23}	λ_1	λ_2	ρ
True	1	-2	1.25	1	-1	0.5	0.4	0.7	0.5
					n=1,000				
Mean	1.011	-2.039	1.272	1.021	-1.022	0.511	0.381	0.693	0.494
Std.dev.	0.197	0.130	0.095	0.152	0.068	0.055	0.062	0.041	0.068
Avg. bias	0.010	-0.037	0.020	0.019	-0.020	0.011	-0.020	-0.008	-0.007
RMSE	0.025	0.009	0.007	0.016	0.005	0.005	0.009	0.003	0.007
I-statistic	1.097	2.121	1.452	1.028	1.129	1.043	2.202	1.153	2.132
P-value	0.546	0.532	0.516	0.473	0.512	0.522	0.527	0.525	0.535
					n=2,000				
Mean	1.009	-2.023	1.261	1.022	-1.018	0.507	0.390	0.694	0.499
Std.dev.	0.137	0.09	0.067	0.106	0.048	0.039	0.044	0.029	0.048
Avg. bias	0.008	-0.021	0.01	0.021	-0.017	0.007	-0.011	-0.007	-0.002
RMSE	0.013	0.005	0.004	0.009	0.003	0.003	0.004	0.002	0.004
I-statistic	1.094	2.013	1.45	1.029	1.119	1.046	2.002	1.149	2.133
P-value	0.502	0.577	0.537	0.499	0.467	0.500	0.546	0.537	0.518

Notes: The I-statistic is calculated following Raftery and Lewis (1992). P-value denotes the p-value of the Geweke (1992) test. The I-statistic and the p-value of the Geweke test are convergence diagnostics of the parameter estimates. The former should be smaller than 5 and the latter should ideally be as large as possible in the interval $(0.1, 1]$ when using customary confidence intervals.

Table 2: Monte Carlo Results: Design 2

	β_{11}	β_{12}	β_{13}	β_{21}	β_{22}	β_{23}	λ_1	λ_2	ρ
True	1	-2	1.25	1	-1	0.5	0.2	0.3	0.5
					n=1,000				
Mean	1.002	-2.032	1.266	1.014	-1.009	0.502	0.162	0.254	0.494
Std. dev.	0.203	0.125	0.088	0.146	0.063	0.050	0.095	0.105	0.065
Avg. bias	0.002	-0.032	0.016	0.014	-0.009	0.002	-0.038	-0.046	-0.006
RMSE	0.033	0.020	0.015	0.024	0.010	0.007	0.015	0.018	0.010
I-statistic	1.140	2.027	1.382	1.038	1.115	1.042	2.560	1.426	2.029
P-value	0.471	0.470	0.452	0.496	0.480	0.483	0.445	0.479	0.454
					n=2,000				
Mean	1.007	-2.017	1.258	1.009	-1.009	0.504	0.188	0.283	0.497
Std. dev.	0.138	0.089	0.066	0.106	0.045	0.038	0.058	0.067	0.045
Avg. bias	0.007	-0.017	0.008	0.009	-0.009	0.004	-0.012	-0.017	-0.003
RMSE	0.023	0.015	0.010	0.017	0.007	0.006	0.009	0.012	0.006
I-statistic	1.102	2.107	1.460	1.035	1.116	1.051	2.263	1.330	2.025
P-value	0.476	0.437	0.441	0.501	0.501	0.470	0.458	0.475	0.443

Notes: The I-statistic is calculated following Raftery and Lewis (1992). P-value denotes the p-value of the Geweke (1992) test. The I-statistic and the p-value of the Geweke test are convergence diagnostics of the parameter estimates. The former should be smaller than 5 and the latter should ideally be as large as possible in the interval $(0.1, 1]$ when using customary confidence intervals.

Table 3: Monte Carlo Results: Design 3

	β_{11}	β_{12}	β_{13}	β_{21}	β_{22}	β_{23}	λ_1	λ_2	ρ
True	1	−2	1.25	1	−1	0.5	0	0.7	0.5
				n=1,000					
Mean	0.965	−2.032	1.268	1.031	−1.025	0.510	−0.111	0.687	0.511
Std. dev.	0.207	0.124	0.087	0.151	0.068	0.052	0.131	0.043	0.067
Avg. bias	−0.035	−0.032	0.018	0.031	−0.025	0.010	−0.111	−0.013	0.011
RMSE	0.032	0.019	0.013	0.026	0.014	0.008	0.038	0.006	0.013
I-statistic	1.186	2.039	1.402	1.044	1.162	1.052	3.133	1.173	2.264
P-value	0.453	0.473	0.465	0.496	0.489	0.491	0.442	0.488	0.451
				n=2,000					
Mean	0.978	−2.013	1.257	1.021	−1.022	0.512	−0.061	0.695	0.514
Std. dev.	0.140	0.088	0.065	0.110	0.048	0.040	0.080	0.028	0.047
Avg. bias	−0.022	−0.013	0.007	0.021	−0.022	0.012	−0.061	−0.005	0.014
RMSE	0.021	0.012	0.010	0.015	0.008	0.006	0.021	0.004	0.008
I-statistic	1.127	2.117	1.454	1.038	1.146	1.058	2.729	1.162	2.232
P-value	0.478	0.437	0.446	0.500	0.501	0.470	0.447	0.482	0.462

Notes: The I-statistic is calculated following Raftery and Lewis (1992). P-value denotes the p-value of the Geweke (1992) test. The I-statistic and the p-value of the Geweke test are convergence diagnostics of the parameter estimates. The former should be smaller than 5 and the latter should ideally be as large as possible in the interval $(0.1, 1]$ when using customary confidence intervals.

Table 4: Monte Carlo Results: Design 4

	β_{11}	β_{12}	β_{13}	β_{21}	β_{22}	β_{23}	λ_1	λ_2	ρ
True	1	−2	1.25	1	−1	0.5	0.4	0.7	0
				n=1,000					
Mean	1.006	−2.035	1.268	1.015	−1.011	0.504	0.376	0.689	−0.003
Std. dev.	0.206	0.131	0.092	0.152	0.068	0.052	0.069	0.043	0.085
Avg. bias	0.006	−0.035	0.018	0.015	−0.011	0.004	−0.024	−0.011	−0.003
RMSE	0.033	0.020	0.014	0.023	0.012	0.009	0.012	0.007	0.012
I-statistic	1.092	1.784	1.280	1.032	1.068	1.036	1.874	1.125	1.864
P-value	0.474	0.466	0.467	0.486	0.474	0.474	0.468	0.476	0.459
				n=2,000					
Mean	1.004	−2.018	1.260	1.005	−1.006	0.503	0.391	0.694	0.000
Std. dev.	0.141	0.094	0.068	0.110	0.048	0.040	0.042	0.029	0.060
Avg. bias	0.004	−0.018	0.010	0.005	−0.006	0.003	−0.009	−0.006	0.000
RMSE	0.022	0.016	0.010	0.016	0.008	0.006	0.007	0.005	0.010
I-statistic	1.071	1.822	1.339	1.031	1.066	1.036	1.660	1.108	1.842
P-value	0.473	0.443	0.450	0.496	0.490	0.483	0.475	0.498	0.461

Notes: The I-statistic is calculated following Raftery and Lewis (1992). P-value denotes the p-value of the Geweke (1992) test. The I-statistic and the p-value of the Geweke test are convergence diagnostics of the parameter estimates. The former should be smaller than 5 and the latter should ideally be as large as possible in the interval $(0.1, 1]$ when using customary confidence intervals.

Table 5: Frequency of firms per industry

Industry	Number of firms	Frequency in %
Food production	212	2.37
Beverage production	53	0.59
Textile industry	647	7.22
Garments and other fiber products	816	9.11
Leather, fur, feathers and related products	153	1.71
Timber processing	116	1.29
Furniture manufacturing	120	1.34
Papermaking and paper products	192	2.14
Printing and record medium production	239	2.67
Cultural, educational and sports goods	263	2.94
Petroleum refining and coking	35	0.39
Raw chemical materials and products	711	7.94
Medical and pharmaceutical products	163	1.82
Chemical fiber products	52	0.58
Rubber products	139	1.55
Plastic products	574	6.41
Nonmetal mineral products	383	4.28
Smelting and pressing of ferrous metals	107	1.19
Smelting and pressing of nonferrous metals	124	1.38
Metal products	854	9.53
Ordinary machinery	779	8.70
Special purpose equipment	436	4.87
Transport equipment	475	5.30
Electronic and telecommunications equipment	736	8.22
Instruments and meters	405	4.52
Artifacts and other manufacturing n.e.c.	175	1.95
Total	8,959	100.00

Table 6: Number of firms in Shanghai

Region	Total	Exporters	Foreign
All areas	8,959	3,064	3,577
Shanghai city district	1,582	573	528
Inner suburbs	4,981	1,557	2,016
Baoshan district	593	143	161
Jiading district	998	271	450
Minhang district	1,175	472	661
Pudong district	2,215	671	744
Outer suburbs	2,199	881	1,005
Fengxian district	521	129	173
Jinshan district	366	156	137
Qingpu district	654	245	320
Songjiang district	658	351	375
Rural area			
Chongming county	197	53	28

Table 7: Frequency of different firm types

		Foreign		
		No	Yes	Total
Exporter	No	4,420	1,475	5,895
	Yes	962	2,102	3,064
	Total	5,382	3,577	8,959

Table 8: Average distance among different firm types (in miles and logs)

Firm type	Average	Std.dev	Maximum
All firms	1.8204	0.0929	2.0209
Domestically-owned exporters	1.8133	0.0941	2.0209
Foreign-owned non-exporters	1.8233	0.0826	2.0209
Foreign-owned exporters	1.8192	0.0868	2.0209
Domestically-owned non-exporters	1.8215	0.0986	2.0209

Table 9: Descriptive statistics covariates

	Total		Exporters		Foreign firms	
	Mean	Std.dev.	Mean	Std.dev.	Mean	Std.dev.
Employment (in logs)	4.695	1.048	5.194	1.086	4.847	1.066
Productivity (in logs)	5.311	0.982	5.306	0.988	5.560	0.986
Intangible asset ratio	0.017	0.046	0.023	0.048	0.022	0.043
Distance to port (in logs)	2.853	0.595	2.906	0.582	2.923	0.558
Distance to city center (in logs)	2.370	0.752	2.365	0.769	2.393	0.679
Average distance to other firms in the neighborhood (in logs)	1.820	0.093	1.817	0.089	1.821	0.085
Number of other firms in the neighborhood (in logs)	7.310	0.859	7.330	0.844	7.393	0.737
Sales to profit ratio	3.011	0.481	3.081	0.455	3.050	0.459
Total assets of smallest exporters (in logs)	7.969	0.763	7.814	0.642	7.918	0.654
Total assets of smallest foreign firms (in logs)	7.967	0.519	7.893	0.454	7.934	0.458

Table 10: Nonspatial and spatial bivariate probit

	Nonspatial model			Spatial model		
	(1)	I-statistic	P-value	(2)	I-statistic	P-value
Dependent variable: exporting indicator (y_{e_i})						
Employment	0.499***	1.001	0.934	0.500***	1.112	0.298
	(0.016)			(0.016)		
Productivity	0.182***	1.112	0.221	0.179***	1.056	0.811
	(0.016)			(0.017)		
Intangible assets ratio	2.151***	0.953	0.853	2.002***	1.059	0.620
	(0.292)			(0.307)		
Distance to port	0.239***	0.953	0.263	0.036	0.950	0.586
	(0.032)			(0.028)		
Distance to city center	0.029	1.001	0.211	0.028	0.953	0.415
	(0.039)			(0.035)		
Average distance to other firms in the neighborhood	−0.746***	1.056	0.140	−0.314*	0.958	0.357
	(0.186)			(0.164)		
Number of other firms in the neighborhood	0.145***	0.953	0.455	0.061**	0.953	0.502
	(0.033)			(0.029)		
Sales to profit ratio	2.177***	0.953	0.403	2.097***	1.056	0.292
	(0.244)			(0.258)		
Sales to profit ratio squared	−0.287***	1.056	0.383	−0.275***	0.952	0.378
	(0.037)			(0.039)		
Total assets of smallest exporters	−0.344***	1.056	0.494	−0.355***	0.953	0.471
	(0.023)			(0.024)		
Constant	−5.430***	1.001	0.800	−4.438***	1.112	0.707
	(0.577)			(0.567)		
Dependent variable: foreign ownership indicator (y_{f_i})						
Employment	0.247***	1.001	0.356	0.242***	1.003	0.421
	(0.014)			(0.015)		
Productivity	0.362***	1.030	0.375	0.353***	0.953	0.501
	(0.016)			(0.017)		
Intangible assets ratio	1.980***	1.056	0.762	1.787***	0.953	0.221
	(0.294)			(0.289)		
Distance to port	0.323***	1.001	0.424	0.037	0.953	0.269
	(0.031)			(0.028)		
Distance to city center	0.426***	1.001	0.541	0.134***	1.112	0.534
	(0.038)			(0.036)		
Average distance to other firms in the neighborhood	−1.431***	0.953	0.603	−0.484***	1.001	0.824
	(0.181)			(0.167)		
Number of other firms in the neighborhood	0.544***	0.953	0.320	0.124***	0.953	0.673
	(0.033)			(0.033)		
Sales to profit ratio	1.922***	0.953	0.535	1.843***	0.952	0.667
	(0.233)			(0.231)		
Sales to profit ratio squared	−0.265***	0.952	0.645	−0.253***	1.056	0.617
	(0.035)			(0.035)		
Total assets of smallest foreign firms	−0.192***	1.056	0.651	−0.187***	1.112	0.409
	(0.029)			(0.030)		
Constant	−8.503***	1.001	0.417	−5.272***	1.001	0.617
	(0.586)			(0.595)		
λ_e				0.733***	1.793	0.296
				(0.035)		
λ_f				0.810***	1.171	0.156
				(0.034)		
ρ	0.610***	1.606	0.923	0.309***	1.001	0.610
	(0.014)			(0.013)		
n	8,959			8,959		

Notes: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Standard errors are in parentheses. Column (1) reports estimates of the nonspatial bivariate probit model. Column (2) reports estimates of the spatial bivariate probit model. The I-statistic is calculated following Raftery and Lewis (1992). P-value denotes the p-value of the Geweke (1992) test. λ_e and λ_f denote the parameter estimates of the spatial lags of the dependent variables, and ρ the tetrachoric correlation. We did 20,000 simulations, of which 4,000 are considered as burn-in. We additionally apply thinning and keep every 10th observation. Thus the estimates are based on 1,600 draws. The significance levels in the table are based on approximations when assuming normally distributed parameter estimates in conjunction with the mean and the standard deviation of each parameter chain. For the parameters based on the Metropolis-Hastings procedure we additionally report the posterior credible 90% intervals here. In the nonspatial model the interval for ρ is $\widehat{\rho} \in [0.586; 0.632]$. The acceptance rate for ρ is 0.595. In the spatial model for the parameters $\{\lambda_e, \lambda_f, \rho\}$, the posterior credible intervals are: $\widehat{\lambda}_e \in [0.674, 0.789]$; $\widehat{\lambda}_f \in [0.754, 0.865]$; $\widehat{\rho} \in [0.287, 0.331]$. The acceptance rates for the same three parameters are 0.472, 0.574, and 0.568, respectively.

Table 11: Robustness: Accounting for state ownership

	Nonspatial model			Spatial model		
	(1)	I-statistic	P-value	(2)	I-statistic	P-value
Dependent variable: exporting indicator (y_{e_i})						
Employment	0.502***	1.056	0.025	0.502***	1.056	0.975
	(0.016)			(0.016)		
Productivity	0.182***	1.001	0.911	0.177***	0.953	0.682
	(0.017)			(0.017)		
Intangible assets ratio	2.151***	0.953	0.359	2.026***	0.953	0.212
	(0.298)			(0.312)		
State-owned	−0.133**	1.001	0.411	−0.157**	0.953	0.131
	(0.057)			(0.061)		
Distance to port	0.236***	0.953	0.461	0.034	1.001	0.514
	(0.032)			(0.029)		
Distance to city center	0.022	0.953	0.963	0.021	0.932	0.582
	(0.040)			(0.035)		
Average distance to other firms in the neighborhood	−0.745***	0.953	0.284	−0.317*	1.112	0.414
	(0.178)			(0.168)		
Number of other firms in the neighborhood	0.143***	0.953	0.531	0.059**	0.953	0.254
	(0.033)			(0.029)		
Total assets of smallest exporters	2.158***	1.001	0.284	2.052***	1.001	0.532
	(0.247)			(0.249)		
Sales to profit ratio	−0.285***	0.953	0.268	−0.269***	1.056	0.519
	(0.038)			(0.038)		
Sales to profit ratio squared	−0.340***	0.953	0.920	−0.352***	1.056	0.354
	(0.023)			(0.025)		
Constant	−5.391***	0.953	0.161	−4.334***	1.056	0.638
	(0.585)			(0.569)		
Dependent variable: foreign ownership indicator (y_{f_i})						
Employment	0.248***	1.001	0.868	0.244***	0.953	0.868
	(0.014)			(0.015)		
Productivity	0.362***	1.001	0.995	0.353***	1.056	0.565
	(0.015)			(0.016)		
Intangible assets ratio	−0.072	0.952	0.620	1.818***	1.001	0.244
	(0.057)			(0.299)		
State-owned	1.994***	1.001	0.607	−0.102*	1.001	0.225
	(0.299)			(0.060)		
Distance to port	0.323***	1.001	0.596	0.036	1.001	0.639
	(0.030)			(0.028)		
Distance to city center	0.423***	0.953	0.659	0.129***	0.953	0.735
	(0.039)			(0.037)		
Average distance to other firms in the neighborhood	−1.436***	0.953	0.452	−0.481***	1.056	0.848
	(0.176)			(0.164)		
Number of other firms in the neighborhood	0.543***	0.953	0.948	0.123***	1.059	0.735
	(0.033)			(0.033)		
Sales to profit ratio	1.906***	1.001	0.295	1.829***	1.001	0.720
	(0.224)			(0.232)		
Sales to profit ratio squared	−0.263***	1.001	0.261	−0.251***	0.952	0.803
	(0.034)			(0.035)		
Total assets of smallest foreign firms	−0.190***	1.056	0.612	−0.185***	0.953	0.782
	(0.029)			(0.030)		
Constant	−8.470***	1.234	0.824	−5.244***	1.001	0.909
	(0.594)			(0.593)		
λ_e				0.740***	1.443	0.649
				(0.036)		
λ_f				0.812***	1.234	0.117
				(0.033)		
ρ	0.610***	1.443	0.911	0.308***	1.112	0.249
	(0.014)			(0.013)		
n	8,959			8,959		

Notes: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Standard errors are in parentheses. Column (1) reports estimates of the nonspatial bivariate probit model. Column (2) reports estimates of the spatial bivariate probit model. The I-statistic is calculated following Raftery and Lewis (1992). P-value denotes the p-value of the Geweke (1992) test. λ_e and λ_f denote the parameter estimates of the spatial lags of the dependent variables, and ρ the tetrachoric correlation. We did 20,000 simulations, of which 4,000 are considered as burn-in. We additionally apply thinning and keep every 10th observation. Thus the estimates are based on 1,600 draws. The significance levels in the table are based on approximations when assuming normally distributed parameter estimates in conjunction with the mean and the standard deviation of each parameter chain. For the parameters based on the Metropolis-Hastings procedure we additionally report the posterior credible 90% intervals here. In the nonspatial model the interval for ρ is $\widehat{\rho} \in [0.588, 0.632]$. The acceptance rate for ρ is 0.506. In the spatial model for the parameters $\{\lambda_e, \lambda_f, \rho\}$, the posterior credible intervals are: $\widehat{\lambda}_e \in [0.674, 0.789]$; $\widehat{\lambda}_f \in [0.754, 0.865]$; $\widehat{\rho} \in [0.287, 0.331]$. The acceptance rates for the same three parameters are 0.472, 0.574, and 0.568, respectively.

Table 12: Robustness: Spatial bivariate probit

	W_{12}			W_8			W_5		
	(1)	I-statistic	P-value	(2)	I-statistic	P-value	(3)	I-statistic	P-value
Dependent variable: exporting indicator (y_{e_i})									
Employment	0.499***	0.979	0.522	0.500***	1.056	0.326	0.498***	1.056	0.337
	(0.016)			(0.016)			(0.016)		
Productivity	0.180***	1.001	0.574	0.178***	1.056	0.768	0.180***	1.056	0.772
	(0.017)			(0.017)			(0.017)		
Intangible assets ratio	2.008***	1.057	0.750	1.990***	1.059	0.556	1.998***	1.059	0.540
	(0.311)			(0.308)			(0.308)		
Distance to port	0.026	1.056	0.215	0.040	0.953	0.981	0.055**	1.001	0.855
	(0.028)			(0.028)			(0.028)		
Distance to city center	0.032	1.001	0.545	0.027	0.953	0.588	0.024	1.056	0.957
	(0.037)			(0.034)			(0.032)		
Average distance to other firms in the neighborhood	−0.380*	1.001	0.445	−0.095	1.056	0.256	0.006	0.953	0.134
	(0.205)			(0.133)			(0.073)		
Number of other firms in the neighborhood	0.062**	0.953	0.713	0.060**	0.953	0.547	0.053**	1.001	0.770
	(0.029)			(0.029)			(0.026)		
Sales to profit ratio	2.070***	0.932	0.905	2.172***	1.056	0.279	2.290***	1.056	0.295
	(0.247)			(0.257)			(0.256)		
Sales to profit ratio squared	−0.272***	0.953	0.950	−0.286***	0.953	0.367	−0.302***	0.953	0.372
	(0.038)			(0.039)			(0.039)		
Total assets of smallest exporters	−0.349***	1.056	0.156	−0.350***	0.953	0.443	−0.351***	0.953	0.372
	(0.024)			(0.024)			(0.024)		
Constant	−4.213***	0.953	0.669	−5.029***	1.112	0.546	−5.361***	1.112	0.621
	(0.590)			(0.552)			(0.534)		
Dependent variable: foreign ownership indicator (y_{f_i})									
Employment	0.242***	1.001	0.599	0.242***	1.003	0.415	0.243***	1.003	0.388
	(0.015)			(0.015)			(0.015)		
Productivity	0.354***	1.001	0.130	0.354***	0.953	0.518	0.356***	0.953	0.539
	(0.016)			(0.017)			(0.017)		
Intangible assets ratio	1.790***	1.112	0.763	1.809***	0.951	0.211	1.816***	0.953	0.209
	(0.304)			(0.289)			(0.289)		
Distance to port	0.026	1.001	0.421	0.042	0.953	0.283	0.070***	0.953	0.187
	(0.028)			(0.027)			(0.027)		
Distance to city center	0.161***	1.112	0.367	0.131***	0.952	0.263	0.140***	0.953	0.323
	(0.038)			(0.035)			(0.033)		
Average distance to other firms in the neighborhood	−0.873***	1.056	0.251	−0.320**	1.056	0.612	0.003	0.953	0.708
	(0.203)			(0.135)			(0.070)		
Number of other firms in the neighborhood	0.131***	1.056	0.695	0.123***	1.055	0.686	0.128***	1.056	0.870
	(0.031)			(0.033)			(0.030)		
Sales to profit ratio	1.895***	0.953	0.559	1.833***	0.952	0.640	1.736***	0.952	0.664
	(0.235)			(0.230)			(0.233)		
Sales to profit ratio squared	−0.261***	1.001	0.566	−0.251***	1.056	0.592	−0.237***	1.056	0.619
	(0.035)			(0.035)			(0.035)		
Total assets of smallest foreign firms	−0.181***	1.056	0.585	−0.208***	1.112	0.296	−0.217***	1.112	0.393
	(0.030)			(0.031)			(0.032)		
Constant	−4.627***	0.953	0.339	−5.449***	1.112	0.920	−5.777***	1.056	0.906
	(0.597)			(0.587)			(0.588)		
λ_e	0.785***	1.443	0.398	0.694***	1.522	0.851	0.619***	1.443	0.937
	(0.037)			(0.034)			(0.030)		
λ_f	0.869***	1.171	0.656	0.768***	1.171	0.565	0.692***	1.001	0.359
	(0.034)			(0.033)			(0.029)		
ρ	0.309***	1.056	0.342	0.309***	1.112	0.448	0.309***	1.171	0.495
	(0.014)			(0.013)			(0.013)		

Notes: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Standard errors are in parentheses. Column (1), (2), and (3) report the results of the spatial bivariate probit model using the weights matrix W_{12}, W_8, and W_5, respectively. The other columns report the I-statistic (Raftery and Lewis, 1992) and the p-value of the Geweke (1992) test. λ_e and λ_f report the parameter estimates of the spatial lags of the dependent variables, and ρ the tetrachoric correlation. We did 20,000 simulations, of which 4,000 are considered as burn-in. We additionally apply thinning and keep every 10th observation. Thus all estimates are based on 1,600 draws. The significance levels in the table are based on approximations when assuming normally distributed parameter estimates in conjunction with the mean and the standard deviation of each parameter chain. For the parameters based on the Metropolis-Hastings procedure we additionally report the posterior credible 90% intervals here. Using W_{12}, for the parameters $\{\lambda_e, \lambda_f, \rho\}$, these are: $\widehat{\lambda}_e \in [0.722, 0.843]$; $\widehat{\lambda}_f \in [0.812, 0.923]$; $\widehat{\rho} \in [0.285, 0.330]$. The acceptance rates for the same three parameters in the Metropolis-Hastings procedure are 0.478, 0.492, and 0.476, respectively. Using W_8, the confidence intervals for $\{\lambda_e, \lambda_f, \rho\}$ are: $\widehat{\lambda}_e \in [0.637, 0.747]$; $\widehat{\lambda}_f \in [0.713, 0.821]$; $\widehat{\rho} \in [0.288, 0.330]$. The acceptance rates for the same three parameters are 0.519, 0.418, and 0.559, respectively. Using W_5, the confidence intervals for $\{\lambda_e, \lambda_f, \rho\}$ are: $\widehat{\lambda}_e \in [0.570, 0.667]$; $\widehat{\lambda}_f \in [0.646, 0.739]$; $\widehat{\rho} \in [0.288, 0.331]$. The acceptance rates for the same three parameters are 0.484, 0.526, and 0.476, respectively.

Table 13: Effect estimates of a one std.dev. increase in explanatory variables on latent export profitability and foreign ownership profitability

			Nonspatial model		Spatial model	
			$\widehat{d}_e^t$	$\widehat{d}_f^t$	$\widehat{d}_e^t$	$\widehat{d}_f^t$
Employment	Min	Mean			0.812	0.459
		Std.err.			0.052	0.046
	p25	Mean			1.772	1.182
		Std.err.			0.225	0.209
	p50	Mean			1.983	1.358
		Std.err.			0.271	0.259
	p75	Mean			2.296	1.615
		Std.err.			0.350	0.324
	Max	Mean			2.894	2.152
		Std.err.			0.492	0.498
	Avg	Mean	0.523	0.259	2.000	1.379
		Std.err.	0.016	0.015	0.279	0.266
Productivity	Min	Mean			0.272	0.627
		Std.err.			0.029	0.060
	p25	Mean			0.592	1.615
		Std.err.			0.089	0.287
	p50	Mean			0.663	1.856
		Std.err.			0.105	0.356
	p75	Mean			0.767	2.209
		Std.err.			0.132	0.445
	Max	Mean			0.967	2.943
		Std.err.			0.181	0.687
	Avg	Mean	0.179	0.355	0.669	1.885
		Std.err.	0.016	0.016	0.108	0.366
Intangible assets ratio	Min	Mean			0.141	0.148
		Std.err.			0.023	0.028
	p25	Mean			0.308	0.381
		Std.err.			0.060	0.093
	p50	Mean			0.345	0.438
		Std.err.			0.069	0.111
	p75	Mean			0.400	0.521
		Std.err.			0.084	0.137
	Max	Mean			0.504	0.694
		Std.err.			0.113	0.201
	Avg	Mean	0.098	0.090	0.348	0.445
		Std.err.	0.013	0.013	0.070	0.114
Distance to port	Min	Mean			0.029	0.034
		Std.err.			0.036	0.044
	p25	Mean			0.070	0.097
		Std.err.			0.056	0.073
	p50	Mean			0.079	0.112
		Std.err.			0.061	0.080
	p75	Mean			0.092	0.134
		Std.err.			0.068	0.093
	Max	Mean			0.117	0.180
		Std.err.			0.081	0.117
	Avg	Mean	0.142	0.192	0.079	0.114
		Std.err.	0.019	0.018	0.061	0.082
Distance to city center	Min	Mean			0.022	0.181
		Std.err.			0.064	0.046
	p25	Mean			0.069	0.463
		Std.err.			0.094	0.127
	p50	Mean			0.080	0.532
		Std.err.			0.100	0.149
	p75	Mean			0.096	0.632
		Std.err.			0.112	0.181
	Max	Mean			0.128	0.840
		Std.err.			0.131	0.253
	Avg	Mean	0.021	0.320	0.081	0.540
		Std.err.	0.029	0.029	0.101	0.152

Continued on next page

Table 13: Effect estimates of a one std.dev. increase in explanatory variables on latent export profitability and foreign ownership profitability

			Nonspatial model		Spatial model	
			$\widehat{d}_e^t$	$\widehat{d}_f^t$	$\widehat{d}_e^t$	$\widehat{d}_f^t$
Average distance to other firms in the neighborhood	Min	Mean			-0.160	-0.376
		Std.err.			0.085	0.142
	p25	Mean			-0.127	-0.283
		Std.err.			0.068	0.103
	p50	Mean			-0.110	-0.238
		Std.err.			0.059	0.086
	p75	Mean			-0.098	-0.207
		Std.err.			0.053	0.074
	Max	Mean			-0.044	-0.081
		Std.err.			0.026	0.027
	Avg	Mean	-0.069	-0.133	-0.111	-0.242
		Std.err.	0.017	0.017	0.059	0.087
Number of other firms in the neighborhood	Min	Mean			0.080	0.191
		Std.err.			0.040	0.045
	p25	Mean			0.176	0.487
		Std.err.			0.085	0.115
	p50	Mean			0.197	0.558
		Std.err.			0.095	0.134
	p75	Mean			0.227	0.664
		Std.err.			0.110	0.161
	Max	Mean			0.287	0.881
		Std.err.			0.139	0.223
	Avg	Mean	0.125	0.467	0.198	0.567
		Std.err.	0.028	0.028	0.096	0.136
Sales to profit ratio	Min	Mean			0.328	0.280
		Std.err.			0.034	0.040
	p25	Mean			0.714	0.722
		Std.err.			0.108	0.150
	p50	Mean			0.799	0.829
		Std.err.			0.128	0.182
	p75	Mean			0.925	0.987
		Std.err.			0.160	0.225
	Max	Mean			1.166	1.315
		Std.err.			0.221	0.338
	Avg	Mean	0.216	0.158	0.806	0.842
		Std.err.	0.018	0.018	0.131	0.187
Total assets of smallest exporters	Min	Mean			-1.495	
		Std.err.			0.270	
	p25	Mean			-1.186	
		Std.err.			0.194	
	p50	Mean			-1.025	
		Std.err.			0.153	
	p75	Mean			-0.915	
		Std.err.			0.129	
	Max	Mean			-0.420	
		Std.err.			0.037	
	Avg	Mean	-0.262		-1.033	
		Std.err.	0.018		0.157	
Total assets of smallest foreign firms	Min	Mean				-0.826
		Std.err.				0.229
	p25	Mean				-0.620
		Std.err.				0.156
	p50	Mean				-0.521
		Std.err.				0.127
	p75	Mean				-0.453
		Std.err.				0.106
	Max	Mean				-0.176
		Std.err.				0.032
	Avg	Mean		-0.099		-0.529
		Std.err.		0.015		0.130

Notes: We report the total effects, $\widehat{d}_e^t$ and $\widehat{d}_f^t$, of one-standard-deviation changes of the regressors on the exporting e and foreign ownership f profitability, respectively. For the nonspatial model, there is only an average effect, while for the spatial model we report the minimum, 25th, 50th, and 75th percentile, maximum, and average for each effect and the corresponding standard error.

Table 14: Effect estimates of a one std.dev. increase in explanatory variables on the marginal probabilities of exporting and foreign ownership

		Nonspatial model		Spatial model	
		$\Delta\widehat{P}_e^t$	$\Delta\widehat{P}_f^t$	$\Delta\widehat{P}_e^t$	$\Delta\widehat{P}_f^t$
Employment	Mean	0.172	0.090	0.540	0.412
	Std.err.	0.006	0.006	0.040	0.054
Productivity	Mean	0.057	0.124	0.209	0.500
	Std.err.	0.006	0.007	0.035	0.045
Intangible asset ratio	Mean	0.031	0.031	0.106	0.146
	Std.err.	0.005	0.005	0.023	0.038
Distance to port	Mean	0.045	0.067	0.023	0.037
	Std.err.	0.006	0.007	0.018	0.027
Distance to city center	Mean	0.007	0.112	0.024	0.177
	Std.err.	0.009	0.011	0.030	0.050
Average distance to other firms in the	Mean	-0.021	-0.045	-0.031	-0.074
neighborhood	Std.err.	0.005	0.006	0.016	0.026
Number of other firms in the neighborhood	Mean	0.040	0.163	0.060	0.186
	Std.err.	0.009	0.010	0.030	0.045
Sales to profit ratio	Mean	0.069	0.055	0.252	0.272
	Std.err.	0.007	0.007	0.041	0.055
Total assets of smallest exporters	Mean	-0.077		-0.231	
	Std.err.	0.006		0.023	
Total assets of smallest foreign firms	Mean		-0.034		-0.154
	Std.err.		0.005		0.033

Notes: We report means and standard errors of changes in the marginal probabilities of exporting e and foreign ownership f resulting from total effects of one-standard-deviation changes of the regressors.

Table 15: Effect estimates of a one std.dev. increase in explanatory variables on the joint probabilities of exporting and foreign ownership

		Nonspatial model			Spatial model		
		$\Delta\widehat{P}_{ef}^t$	$\Delta\widehat{P}_{e0}^t$	$\Delta\widehat{P}_{0f}^t$	$\Delta\widehat{P}_{ef}^t$	$\Delta\widehat{P}_{e0}^t$	$\Delta\widehat{P}_{0f}^t$
Employment	Mean	0.125	0.047	-0.035	0.499	0.041	-0.087
	Std.err.	0.006	0.005	0.004	0.055	0.045	0.029
Productivity	Mean	0.077	-0.020	0.047	0.285	-0.076	0.216
	Std.err.	0.005	0.003	0.005	0.036	0.016	0.040
Intangible asset ratio	Mean	0.028	0.003	0.004	0.091	0.015	0.055
	Std.err.	0.004	0.003	0.003	0.018	0.016	0.026
Distance to port	Mean	0.049	-0.004	0.018	0.019	0.004	0.018
	Std.err.	0.005	0.004	0.005	0.013	0.012	0.019
Distance to city center	Mean	0.041	-0.035	0.071	0.056	-0.032	0.121
	Std.err.	0.008	0.004	0.008	0.023	0.016	0.039
Average distance to other firms in the	Mean	-0.027	0.006	-0.018	-0.052	0.021	-0.023
neighborhood	Std.err.	0.004	0.003	0.004	0.016	0.013	0.015
Number of other firms in the neighborhood	Mean	0.078	-0.039	0.085	0.079	-0.020	0.107
	Std.err.	0.008	0.004	0.008	0.023	0.017	0.034
Sales to profit ratio	Mean	0.057	0.012	-0.002	0.222	0.030	0.050
	Std.err.	0.005	0.004	0.005	0.038	0.029	0.036
Total assets of smallest exporters	Mean	-0.042	-0.035	0.042	-0.150	-0.081	0.150
	Std.err.	0.003	0.003	0.003	0.016	0.007	0.016
Total assets of smallest foreign firms	Mean	-0.013	0.013	-0.021	-0.078	0.078	-0.076
	Std.err.	0.002	0.002	0.003	0.018	0.018	0.015

Notes: We report means and standard errors of changes in the joint probabilities of exporting e and foreign ownership f, e.g. ef, $e0$, and $0f$, resulting from total effects of one-standard-deviation changes of the regressors.

Table 16: A more general spatial bivariate probit

	(1)	I-statistic	P-value
Dependent variable: exporting indicator (y_{e_i})			
Employment	0.493***	1.062	0.921
	(0.016)		
Productivity	0.173***	1.062	0.459
	(0.018)		
Intangible assets ratio	1.941***	1.063	0.608
	(0.301)		
Distance to port	0.029	1.060	0.371
	(0.029)		
Distance to city center	0.025	1.238	0.183
	(0.046)		
Average distance to other firms in the neighborhood	−0.279*	0.985	0.858
	(0.168)		
Number of other firms in the neighborhood	0.053	1.445	0.193
	(0.046)		
Sales to profit ratio	2.107***	1.062	0.535
	(0.246)		
Sales to profit ratio squared	−0.278***	1.062	0.524
	(0.037)		
Total assets of smallest exporters	−0.336***	1.147	0.321
	(0.023)		
Constant	−4.473***	1.062	0.592
	(0.616)		
Dependent variable: foreign ownership indicator (y_{f_i})			
Employment	0.238***	0.985	0.180
	(0.015)		
Productivity	0.350***	0.985	0.997
	(0.016)		
Intangible assets ratio	1.764***	1.062	0.371
	(0.302)		
Distance to port	0.022	1.238	0.180
	(0.029)		
Distance to city center	0.155***	1.337	0.158
	(0.043)		
Average distance to other firms in the neighborhood	−0.490***	0.985	0.704
	(0.164)		
Number of other firms in the neighborhood	0.140***	1.337	0.280
	(0.041)		
Sales to profit ratio	1.832***	0.985	0.286
	(0.229)		
Sales to profit ratio squared	−0.252***	0.985	0.373
	(0.035)		
Total assets of smallest foreign firms	−0.177***	0.985	0.254
	(0.029)		
Constant	−5.349***	1.062	0.591
	(0.600)		
λ_{ee}	0.760***	3.932	0.112
	(0.071)		
λ_{ef}	0.013	1.147	0.164
	(0.078)		
λ_{fe}	0.129	2.739	0.142
	(0.081)		
λ_{ff}	0.752***	2.177	0.288
	(0.067)		
ρ	0.593***	1.337	0.537
	(0.015)		
n	8,959		

Notes: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Standard errors are in parentheses. Column (1) reports estimates of the more general spatial bivariate probit model. The I-statistic is calculated following Raftery and Lewis (1992). P-value denotes the p-value of the Geweke (1992) test. λ_{ee}, λ_{ef}, λ_{fe}, and λ_{ff} denote the parameter estimates of the spatial lags of the dependent variables, and ρ the tetrachoric correlation. We did 20,000 simulations, of which 7,500 are considered as burn-in. We additionally apply thinning and keep every 12th observation. The significance levels in the table are based on approximations when assuming normally distributed parameter estimates in conjunction with the mean and the standard deviation of each parameter chain. For the parameters based on the Metropolis-Hastings procedure we additionally report the posterior credible 90% intervals here. For the parameters $\{\lambda_{ee}, \lambda_{ef}, \lambda_{fe}, \lambda_{ff}, \rho\}$, these are: $\widehat{\lambda}_{ee} \in [0.633, 0.865]$; $\widehat{\lambda}_{ef} \in [-0.102, 0.157]$; $\widehat{\lambda}_{fe} \in [-0.020, 0.259]$; $\widehat{\lambda}_{ff} \in [0.644, 0.868]$; $\widehat{\rho} \in [0.568, 0.616]$. The acceptance rates for the same parameters in the Metropolis-Hastings procedure are 0.465,0.440,0.473,0.448, and 0.584, respectively.

Table 17: Nonspatial and spatial univariate probit

	Nonspatial model			Spatial model		
	(1)	I-statistic	P-value	(2)	I-statistic	P-value
Dependent variable: exporting indicator (y_{e_i})						
Employment	0.499***	0.953	0.684	0.493***	1.112	0.192
	(0.016)			(0.016)		
Productivity	0.169***	0.953	0.438	0.179***	1.001	0.630
	(0.016)			(0.017)		
Intangible assets ratio	2.010***	1.112	0.685	1.983***	1.001	0.234
	(0.301)			(0.302)		
Distance to port	0.254***	1.056	0.255	0.028	1.001	0.966
	(0.031)			(0.029)		
Distance to city center	0.007	1.001	0.557	0.028	1.056	0.154
	(0.039)			(0.034)		
Average distance to other firms in the neighborhood	−0.682***	1.112	0.404	−0.299*	0.953	0.563
	(0.178)			(0.167)		
Number of other firms in the neighborhood	0.146***	1.001	0.336	0.058**	1.171	0.845
	(0.032)			(0.028)		
Sales to profit ratio	2.117***	0.953	0.577	2.036***	0.953	0.670
	(0.241)			(0.248)		
Sales to profit ratio squared	−0.277***	1.001	0.613	−0.265***	0.953	0.614
	(0.037)			(0.037)		
Total assets of smallest exporters	−0.384***	1.001	0.978	−0.363***	1.056	0.380
	(0.033)			(0.025)		
Constant	−5.066***	1.001	0.860	−4.223***	0.953	0.687
	(0.619)			(0.564)		
λ_e				0.751	2.008	0.237
				(0.039)		
Dependent variable: foreign ownership indicator (y_{f_i})						
Employment	0.249***	0.953	0.637	0.241***	1.056	0.965
	(0.015)			(0.015)		
Productivity	0.362***	1.001	0.672	0.347***	1.056	0.404
	(0.016)			(0.017)		
Intangible assets ratio	1.935***	1.171	0.705	1.751***	0.932	0.476
	(0.292)			(0.286)		
Distance to port	0.320***	0.953	0.113	0.032	1.056	0.169
	(0.031)			(0.029)		
Distance to city center	0.429***	1.001	0.479	0.129***	1.056	0.410
	(0.039)			(0.035)		
Average distance to other firms in the neighborhood	−1.435***	1.112	0.491	−0.457***	1.056	0.479
	(0.174)			(0.162)		
Number of other firms in the neighborhood	0.545***	0.953	0.336	0.117***	1.112	0.353
	(0.032)			(0.032)		
Sales to profit ratio	1.887***	1.001	0.695	1.810***	1.001	0.593
	(0.222)			(0.231)		
Sales to profit ratio squared	−0.259***	1.001	0.712	−0.248***	1.001	0.448
	(0.034)			(0.035)		
Total assets of smallest foreign firms	−0.201***	0.953	0.842	−0.187***	1.003	0.382
	(0.030)			(0.031)		
Constant	−8.391***	1.001	0.985	−5.159***	1.001	0.784
	(0.595)			(0.578)		
λ_f				0.823***	1.443	0.571
				(0.033)		
n	8,959			8,959		

Notes: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Standard errors are in parentheses. Column (1) reports estimates of the nonspatial probit model. Column (2) reports estimates of the spatial probit model. The I-statistic is calculated following Raftery and Lewis (1992). P-value denotes the p-value of the Geweke (1992) test. λ_e and λ_f denote the parameter estimates of the spatial lags of the dependent variables. We did 10,000 simulations, of which 2,000 are considered as burn-in. We additionally apply thinning and keep every 5th observation. The significance levels in the table are based on approximations when assuming normally distributed parameter estimates in conjunction with the mean and the standard deviation of each parameter chain. For the parameters based on the Metropolis-Hastings procedure we additionally report the posterior credible 90% intervals. For the parameters $\{\lambda_e, \lambda_f\}$, these are: $\widehat{\lambda}_e \in [0.685, 0.813]$ and $\widehat{\lambda}_f \in [0.768, 0.878]$. The acceptance rates for the same parameters in the Metropolis-Hastings procedure are 0.465 and 0.441, respectively.

Appendix

Appendix A: A more general spatial bivariate probit model

In this Appendix, we discuss a natural extension of the model, where the vector of latent outcomes y_h^* does not only depend on its own spatial lag, $\overline{y}_h^*$, but also the one of the other outcome, $\overline{y}_g^*$ with $g \neq h$. The correspondingly modified model could be written as

$$y_h^* = \lambda_{hh}\overline{y}_h^* + \lambda_{hg}\overline{y}_g^* + x_h\beta_h + u_h, \tag{12}$$

After defining $\Lambda = \begin{pmatrix} \lambda_{ee} & \lambda_{ef} \\ \lambda_{fe} & \lambda_{ff} \end{pmatrix}$, whose typical element we will refer to as λ_{hg}, we can write the stacked counterpart to equation (7) for both equations together as

$$\begin{pmatrix} y_e^* \\ y_f^* \end{pmatrix} = (\Lambda \otimes W) \begin{pmatrix} y_e^* \\ y_f^* \end{pmatrix} + \begin{pmatrix} x_e & 0 \\ 0 & x_f \end{pmatrix} \begin{pmatrix} \beta_e \\ \beta_f \end{pmatrix} + \begin{pmatrix} u_e \\ u_f \end{pmatrix}, \tag{13}$$

where the error process is assumed to be the same as in the main text.

Priors

Using the same priors as before and assuming

$$\lambda_{hg} \sim U(-1, 1)$$

for all $\{hg\}$, it is straightforward to extend the more restrictive set-up in the main text when using the likelihood function below.

Likelihood

When introducing the notation of

$$\begin{aligned} y^* &= (y_e^{*\prime}, y_f^{*\prime})' \\ L &= (I_{2n} - \Lambda \otimes W) = (L_{gh}), \\ \widetilde{L} &= L^{-1} = (\widetilde{L}_{gh}), \\ X &= diag_{h\in\{e,f\}}(x_h) \\ \beta &= (\beta_e', \beta_f')' \\ u &= (u_e', u_f')', \end{aligned}$$

the reduced form of the latent process can be written as

$$y^* = \widetilde{L}X\beta + \widetilde{L}u. \tag{14}$$

Then, we can write the likelihood of the joint distribution of (y_e^*, y_f^*) as

$$p(y^*|\beta, \Lambda, \rho, X, W) = \frac{1}{2\pi^n|\Sigma|^{n/2}}|L_{ff}||L_{ee} - L_{ef}L_{ff}^{-1}L_{fe}|exp\left[-\frac{1}{2}trace\left(R\Sigma^{-1}\right)\right],$$

where, under the present assumptions, Σ is the same as in the main text and $R = \begin{pmatrix} r_{ee} & r_{ef} \\ r_{fe} & r_{ff} \end{pmatrix}$ is a 2×2 matrix containing the elements

$$r_{ee} = (L_{ee}y_e^* + L_{ef}y_f^* - x_e\beta_e)'(L_{ee}y_e^* + L_{ef}y_f^* - x_e\beta_e) \tag{15}$$

$$r_{ef} = (L_{ee}y_e^* + L_{ef}y_f^* - x_e\beta_e)'(L_{fe}y_e^* + L_{ff}y_f^* - x_f\beta_f) \tag{16}$$

$$r_{fe} = (L_{fe}y_e^* + L_{ff}y_f^* - x_f\beta_f)'(L_{ee}y_e^* + L_{ef}y_f^* - x_e\beta_e) \tag{17}$$

$$r_{ff} = (L_{fe}y_e^* + L_{ff}y_f^* - x_f\beta_f)'(L_{fe}y_e^* + L_{ff}y_f^* - x_f\beta_f). \tag{18}$$

The joint distribution of y^* is

$$y^* \sim N(\widetilde{L}X\beta, \widetilde{L}(\Sigma \otimes I_n)\widetilde{L}').$$

Define the precision matrix H, which is the inverse of the variance of y^*,

$$H = (\widetilde{L}\Psi\widetilde{L}')^{-1} = L'(\Sigma^{-1} \otimes I_n)L = \begin{pmatrix} H_{ee} & H_{ef} \\ H_{fe} & H_{ff} \end{pmatrix}.$$

Then, the conditional distributions of y_e^* and y_f^* can be formulated using the variance of y^* or the precision matrix H (see Geweke, 2005) as

$$y_e^*|\theta_{-y_e^*} \sim N\bigg(\widetilde{L}_{ee}x_e\beta_e + \widetilde{L}_{ef}x_f\beta_f - H_{ee}^{-1}H_{ef}(y_f - \widetilde{L}_{fe}x_e\beta_e - \widetilde{L}_{ff}x_f\beta_f), H_{ee}^{-1}\bigg),$$

$$y_f^*|\theta_{-y_f^*} \sim N\bigg(\widetilde{L}_{fe}x_e\beta_e + \widetilde{L}_{ff}x_f\beta_f - H_{ff}^{-1}H_{fe}(y_e - \widetilde{L}_{ee}x_e\beta_e - \widetilde{L}_{ef}x_f\beta_f), H_{ff}^{-1}\bigg).$$

The conditional distribution of $\beta = (\beta_e', \beta_f')'$ is

$$\beta|\theta_{-\beta} \propto N(\overline{\beta}, \overline{V}_\beta),$$

where

$$\overline{\beta} = \overline{V}_{\beta}\left(X'\left(\Sigma^{-1}\otimes I_n\right)Ly^* + \underline{V}^{-1}\underline{\delta}\right),$$
$$\overline{V}_{\beta} = \left(X'\left(\Sigma^{-1}\otimes I_n\right)X + \underline{V}^{-1}\right)^{-1}.$$

Notice that we consider drawing the elements of β jointly, here. In the main text, we outlined the procedure for drawing them separately for both equations. Either approach is applicable.

The conditional distributions of λ_{ee},λ_{ef},λ_{fe}, and λ_{ff} are

$$\lambda_{ee}|\theta_{-\lambda_{e}e} \propto |L_{ee} - L_{ef}L_{ff}^{-1}L_{fe}|exp\left[-\frac{1}{2}trace\left(R\Sigma^{-1}\right)\right] \quad (19)$$

$$\lambda_{ef}|\theta_{-\lambda_{e}f} \propto |L_{ee} - L_{ef}L_{ff}^{-1}L_{fe}|exp\left[-\frac{1}{2}trace\left(R\Sigma^{-1}\right)\right] \quad (20)$$

$$\lambda_{fe}|\theta_{-\lambda_{f}e} \propto |L_{ee} - L_{ef}L_{ff}^{-1}L_{fe}|exp\left[-\frac{1}{2}trace\left(R\Sigma^{-1}\right)\right] \quad (21)$$

$$\lambda_{ff}|\theta_{-\lambda_{f}f} \propto |L_{ff}||L_{ee} - L_{ef}L_{ff}^{-1}L_{fe}|exp\left[-\frac{1}{2}trace\left(R\Sigma^{-1}\right)\right] \quad (22)$$

The conditional distribution of ρ reads exactly as in equation (9), except that the definition of R is different here from the one in the main text.

Modified application

We estimate the model outlined above on the data for Shanghai as used in the main text. While a single draw of the Monte Carlo chain (of which there were 20,000) took approximately 1.7 seconds in the baseline specification in the main text on average, a draw with the more complex model outlined in Appendix A takes approximately 5.9 seconds. The results are summarized in Table 16, and they suggest that the data do not support spillovers of the profitability of foreign ownership on exporting and vice versa (this can be seen from the statistically insignificant parameters ($\widehat{\lambda}_{ef}, \widehat{\lambda}_{fe}$) in the table). Given this, we refrain from an in-depth discussion of the corresponding results.

– Table 16 about here –

Appendix B: Further extensions and computational issues

In principal, it is possible to take the model proposed in this paper to problems beyond two structural equations. Such a multivariate probit model would be relatively straightforward to analyze as long as the spatial-lag-parameter-matrix Λ as introduced in the Appendix is diagonal, which was the case in the main text. For non-diagonal Λ, we saw that the computational burden was already significant with only two equations. Hence, higher-dimensional and more general model versions would call for fast approximation algorithms of determinants and inverses to be applicably with large data-sets as the one used in this paper.

Appendix C: Results for univariate spatial and nonspatial probit models

In Table 17, we present the results of nonspatial and spatial univariate probit models. These models correspond to the ones estimated in Table 10, except that $\rho = 0$.

– Table 17 about here –

The numbers in Table 17 suggest that there are smaller differences between the point estimates and the standard errors between the equations in Table 17 and their counterparts in Table 10. In any case, ignoring that $\rho \neq 0$ appears to entail less of a problem with the data at hand than when ignoring the spatial lags of the latent dependent variables in the model specification.