On the use of network matrices with Sherman-Morrison-Woodbury inverses in econometrics

Peter H. Egger*
ETH Zurich
CEPR, CESifo, WIFO

Kevin E. Staub†
University of Melbourne
IZA

August 19, 2025

Abstract

We discuss and motivate the use of specific network matrices with a Sherman-Morrison-Woodbury inverse form in econometrics. Such matrices feature a parsimonious structure dependent on only two vectors that fits many situations where link intensity is the product of two size-based measures — a common reduced-form approximation in many gravity and industrial organization contexts. Matrices of Sherman-Morrison-Woodbury form permit special treatment which is fundamentally different from customary network matrices. In particular, the use of such matrices can lead to network models where the key slope parameters can be estimated easily by OLS from the structural network model. We illustrate their properties theoretically and in simulations.

Keywords: Network econometrics; Sherman-Morrison-Woodbury inverses.

JEL-codes: C13; C18; C23; C31; C33; F14.

*ETH Zurich, Department of Management, Technology, and Economics, Weinbergstr. 35, 8092 Zurich, Switzerland; E-mail: egger@kof.ethz.ch; + 41 44 632 41 08.

†Department of Economics, 111 Barry Street, The University of Melbourne, 3010 VIC, Australia; E-mail: kevin.staub@unimelb.edu.au; +61 3 903 53776.

1 Introduction

Spatial and network econometrics have both seen a vibrant development and growing interest for several decades. Harry H. Kelejian and his co-authors had spurred this development and interest with fundamental contributions by improving our understanding about estimation, identification, as well as testing (see Kelejian and Prucha, 1997, 1998, 1999, 2001, 2004, 2007, 2010, Kelejian, 2008, Kelejian et al., 2004, Kapoor et al., 2007, Kelejian and Piras, 2011, 2014, 2016, 2017, 2018, 2020, Kelejian et al., 2004, and Kelejian and Robinson, 1992, 1998, 2004). Among the many mentioned contributions, we would like to highlight those that proposed methods permitting working with large cross sections of units – in cross-section or panel-data environments – with large networks of flexible form.

In this paper, we focus on the use of network matrices with a peculiar form in the estimation of models that are linear in parameters. These matrices differ from conventional network matrices in that they do not have zero diagonal entries – i.e., they feature what is called *self-reflection* – and their entries can be characterized by just two vectors of real elements. Network matrices of the kind considered here lead to a Leontief-type inverse matrix of the Sherman-Morrison-Woodbury form (see Woodbury, 1949, 1950, Sherman and Morrison, 1950, Bartlett, 1951, Riedel, 1992, and Hao and Simoncini, 2021). This form brings about special considerations for estimation and the identifiability of parameters. Despite having the form of a conventional SAR (network or spatial autoregressive) model, the key slope parameters can be estimated consistently by linear OLS regressions without needing to resort to IV (instrumental variable) estimation of the structural model equation or nonlinear estimation of its reduced form. The reason is that the reduced form of the model is itself linear. However, depending on the specific configuration and properties of the Sherman-Morrison-Woodbury network matrix, only some parameters might be identified.

Network matrices with a Sherman-Morrison-Woodbury inverse may arise in a number of models in economics and finance, including in the context of international trade flows, regional economics and migration flows, supply chains, and banking and financial linkages. In the paper, to show the practical relevance of models with such network matrices, we focus on an example from the context of log-linearized structural gravity-type flow equations.

The remainder of the paper is organized as follows. The next section introduces the concept of Sherman-Morrison-Woodbury inverses. We outline a generic structure that is general enough to be able to cover both cross-section and panel data cases. In Section 3 we specifically focus on panel data and discuss an application to a model of international trade flows. In Section 4, we provide finite-sample-performance evidence of cross-section and panel data estimators based on Monte Carlos simulations. We conclude with a short summary in the last section.

2 Models of interest involving network matrices with Leontief-type Sherman-Morrison-Woodbury inverses

Let us introduce an observation index $j = 1, ..., n$ with n being the sample size, noting that in cross-sectional data $j = i$ with $i = 1, ..., N$ and $n = N$, and with cross-section×time panel data $j = \{it\}$ with $i = 1, ..., N$, $t = 1, ..., T$, and $n = NT$. Let y_j denote a scalar-valued stochastic outcome variable, x_j be a $1 \times K$ vector of nonstochastic regressors. Let $w_{jj'}$ and $m_{jj'}$ be network weights measuring the known relative importance of unit j' for unit j, and let u_j be a stochastic term. Then, denoting the unknown parameters by Greek letters, the model we consider in this paper is

$$y_j = \lambda \sum_{j'=1}^{n} w_{jj'} y_{j'} + \alpha + x_j \beta + u_j, \quad u_j = \rho \sum_{j'=1}^{n} m_{jj'} u_{j'} + e_j, \tag{1}$$

where λ, α and ρ are scalars, and β is a $K \times 1$ vector. The key objects of interest we focus on in this paper are the slope parameters in the systematic part of the model, λ and β, which we will also call the structural slopes. Stacking the data, using ι_n for an $n \times 1$ vector of ones, and collecting the elements $w_{jj'}$ and $m_{jj'}$ in the $n \times n$ matrices W and M, respectively, we obtain the structural and reduced model forms as

$$y = \lambda W y + \alpha \iota_n + X\beta + u, \quad u = \rho M u + e, \tag{2}$$

$$y = (I_n - \lambda W)^{-1}(\alpha \iota_n + X\beta + u), \quad u = (I_n - \rho M)^{-1} e. \tag{3}$$

Customary properties of network matrices $V = \{W, M\}$: Customary properties of a network matrix V in econometrics are the following: (i) its elements are all real valued and typically assumed to be nonnegative (although the latter is not key); (ii) its diagonal values are all zero to avoid self-reflection; (iii) the elements $v_{jj'}$ are absolutely summable which can be achieved by row normalization or scalar normalization (such as maximum row- or column-sum normalization; see Kelejian and Prucha, 2010; in practice, row normalization enjoys the most frequent use of those).

Sherman-Morrison-Woodbury forms of V: In this case, the network matrix is constructed based on two $n \times 1$ vectors $v_{1,n}$ and $v_{2,n}$ as

$$V = v_{1,n} v'_{2,n}, \qquad v_{1,n}, v_{2,n} \in \mathbb{R}^n. \tag{4}$$

It should be noted that, as long as the signs of all elements in $v'_{2,n}$ are the same, we can always normalize it such that its elements sum up to unity and scale $v_{1,n}$ with the row sum of the original $v'_{2,n}$. Clearly, pre-multiplying any vector of matrix by $v'_{2,n}$ then produces weighted average elements of that vector or matrix in a row vector.

Economic interpretation and examples: Network matrices of the form (4) are parsimonious relative to conventional ones in that they only specify $2n$ elements instead of $n(n-1)$. They also are easily interpretable in terms of offering clear 'sender' and 'receiver' channels (or out-strength and in-strength in network terms). This structure fits many situations where link intensity is the product of two size-based measures — a common reduced-form approximation in many gravity and industrial organization contexts. E.g., suppose that $v_{1,n}$ is a vector of sales shares or sales values, and $v_{2,n}$ is a vector of purchase shares or purchase values. Then, the matrix V would contain higher-valued entries for entities (firms, regions, countries) with higher products of sales and purchase values. The same would be true for population shares or counts; out- and in-citations shares of counts regarding patents or academic output; shares of a region's out- and in-migration flows; banks' shares of total interbank lending and borrowing; etc. However, Sherman-Morrison-Woodbury forms of network matrices also can arise from structural economic models (see the next section for an example from the trade literature).

Inverse of Sherman-Morrison-Woodbury matrices: Let $v^\circ_{1,n} \in \mathbb{R}^n$ and $v^\circ_{2,n} \in \mathbb{R}^n$ be generic column vectors of n raw elements each, and let ι_n be an $n \times 1$ vector of ones. We will consider the case, where all elements of $v^\circ_{1,n}$ have the same sign, and all elements of $v^\circ_{2,n}$ have the same sign. It is probably fair to say that typically one would encounter cases, where all elements of these vectors are positive (at least nonnegative). Then, $v^{\circ\prime}_{2,n}\iota_n = v^\circ_{2,sum} \in \mathbb{R}$. We will generally normalize $v_{2,n} = (v^\circ_{2,sum})^{-1}v^\circ_{2,n}$, whereby $v'_{2,n}\iota_n = 1$ for $v \in \{w, m\}$. Moreover, we will normalize $v^\circ_{1,n}$ by its maximum element. Defining $v^\circ_{1,max} = max_{i,\ldots,n}(v^\circ_{1,n,i})$, we will obtain $v_{1,n} = (v^\circ_{1,max})^{-1}v^\circ_{1,n}$. Let us collect $(v^\circ)^{-1} = (v^\circ_{1,max}v^\circ_{2,sum})^{-1}$. Then, we could state $V^\circ = v^\circ_{1,n}v^{\circ\prime}_{2,n} = (v^\circ)^{-1}V$. With this normalization, the maximum row sum of V would be unity under the adopted assumptions. Let us further introduce real-valued parameter scalars $\theta_{V,0}$ and collect terms into the real-valued $\theta_V = \theta_{V,0}(v^\circ)^{-1}$.

The Sherman-Morrison-Woodbury inverse of a matrix $(I_n - \theta_V V) = (I_n - \theta_V v_{1,n}v'_{2,n})$ with $\theta_V \in \mathbb{R}$ and $v_{1,n}, v_{2,n} \in \mathbb{R}^n$ is defined as

$$(I_n - \theta_V V)^{-1} \quad = \quad I_n + \frac{\theta_V}{1 - \theta_V v'_{2,n}v_{1,n}}V = I_n + \frac{\theta_V}{1 - \theta_V \omega_V}V, \tag{5}$$

which obviously requires $\omega_V = v'_{2,n}v_{1,n} \neq \theta_V^{-1}$. In our case, $\theta_V V \in \{\lambda W, \rho M\}$.

Upon defining the *ex ante* known scalar $\omega_V = v'_{2,n}v_{1,n}$ and the auxiliary parameters $\lambda^* = \frac{\lambda}{1-\lambda\omega_W}$ and $\rho^* = \frac{\rho}{1-\rho\omega_M}$ with $\{\lambda^*, \rho^*\} \neq \{\omega_W^{-1}, \omega_M^{-1}\}$, we obtain

$$(I_n - \lambda W)^{-1} = I_n + \lambda^* W, \quad (I_n - \rho W)^{-1} = I_n + \rho^* M \tag{6}$$

as well as

$$(I_n - \lambda W)^{-1}(I_n - \rho M)^{-1} = I_n + \lambda^* W + \rho^* M + \lambda^* \rho^* W M. \tag{7}$$

For any conformable variable or vector h, let $Vh = \kappa_{V,h} v_{1,n}$, where $\kappa_{V,h}$ is a scalar, and $V \in \{W, M\}$ as well as $v \in \{w, m\}$. Accordingly, $u = (\rho^* \kappa_{M,e} m_{1,n} + e)$.

Let us define the scalars $v'_{2,n} y = \kappa_{V,y}$ and $v'_{2,n} e = \kappa_{V,e}$ for $v \in \{w, m\}$ and $V \in \{W, M\}$. Then, note that

$$Wy = \kappa_{W,y} w_{1,n}, \qquad Me = \kappa_{M,e} m_{1,n}. \tag{8}$$

Obviously, for any column h of the matrix (ι_n, X), $Vh = \kappa_{V,h} v_{1,n}$. Hence, $w_{1,n}$ and $\kappa_{W,h} w_{1,n}$, $\kappa_{W,h'} w_{1,n}$, as well as $\kappa_{W,y} w_{1,n}$ are pairwise perfectly collinear.

Reconsidering the structural and reduced-form models: With the above definitions and notions at hand, we can restate the structural and reduced network-model forms involving network matrices of the Sherman-Morrison-Woodbury form as

$$y = \lambda \kappa_{W,y} w_{1,n} + \alpha \iota_n + X\beta + \rho^* \kappa_{M,e} m_{1,n} + e, \tag{9}$$

$$y = (I_n + \lambda^* W)[\alpha \iota_n + X\beta + (\rho^* \kappa_{M,e} m_{1,n} + e)]. \tag{10}$$

Let us now use the short hand

$$\lambda^* \kappa_Z w_{1,n} = \lambda^* W[\alpha \iota_n + X\beta + (\rho^* \kappa_{M,e} m_{1,n} + e)]. \tag{11}$$

This permits expressing the reduced-form model more compactly as

$$y = \alpha \iota_n + X\beta + \rho^* \kappa_{M,e} m_{1,n} + \lambda^* \kappa_Z w_{1,n} + e. \tag{12}$$

Parameter identification and estimation: What stands out with the structural and reduced-form models using Sherman-Morrison-Woodbury-type network matrices is that in both cases the model error is the same $(\rho^* \kappa_{M,e} m_{1,n} + e)$ and that this error is mean independent of the regressors if $E(m_{1,n}|w_{1,n}, X) = 0$, a condition that can be verified empirically as it involves known variable vectors. Thus, assuming m_{1n} is exogenous and with the residuals e being IID, (i) OLS regression of y on Wy and X (that is, of the structural model, eq. (9)) can estimate the structural slope parameters λ and β consistently; (ii) an even simpler OLS regression of y on $w_{1,n}$ and X estimates β consistently. In such a regression, $w_{1,n}$ controls for network interdependence in the outcome and its coefficient captures $\lambda \kappa_{W,y}$ (structural model), or, equivalently, $\lambda^* \kappa_Z$ (reduced-form model, eq. (12)). For both β and λ to be identified by the OLS approach (i), the network matrices W and M need to satisfy certain conditions, which we summarize in Table 1. To see this, let v_2 be generally be a vector of shares that sum up to unity for $v \in \{w, m\}$. Then, everything depends on v_1. If all elements of v_1 are identical, we can call this a case of row normalization. In that case, $v_1 = v_0 \iota_N$, where v_0 is a real-valued scalar and ι_N is a column vector of ones with N rows. With row normalization, any variable multiplied by V is a manifold of ι_N. Therefore, λ could not be identified in that case. If the elements of v_1 vary across the rows, this is analogous to the case of scalar normalization of $V \in \{W, M\}$. This permits an identification of λ as long as $W \neq M$ (or ρM is absent from the model).

Table 1: STRUCTURAL SLOPE PARAMETER IDENTIFICATION WITH DIFFERENT NETWORK MATRIX CONFIGURATIONS

Configuration of network matrices $V \in \{W, M\}$	Identified	Not identified
Scalar-normalized (W, M) with $W \neq M$	λ, β	—
Scalar-normalized $W = M$	β	λ
Scalar-normalized W, row-normalized M	λ, β	—
Row-normalized W and M	β	λ

A further and related identification issue arises from the fact mentioned above that pre-multiplying the regressors by W leads to perfect multicollinearity (columns of WX are pairwise collinear). Thus, a specification of the model with contextual effects in in the sense of an inclusion of WX (with parameter vector, say, γ) in the structural model is not identified. Hence, the Sherman-Morrison-Woodbury network form precludes the identification of parameters (λ, γ) as included in spatial or network Durbin-type models (see Debarsy, 2012), in both their structural and reduced forms.

Variance-covariance matrix of disturbances: In the light of the assumptions of $E(e) = 0$ $E(ee') = \sigma_e^2 I_n$, the above expressions, and the definition of the regression residual of the structural model OLS estimation, $\hat{u} = y - \hat{\lambda} W y + \hat{\alpha} - X\hat{\beta}$, we can state the following. Let $Z = (\iota_n, Wy, X)$ and $\theta = (\alpha, \lambda, \beta')'$ with $\dim(\theta){=}K$. The OLS variance estimator for the coefficients of the structural model is $\hat{V}(\hat{\theta}) = s_u^2 Z'Z$, where $s_u^2 \hat{u}'\hat{u}/(N-K)$. This estimator is consistent if m_{1n} is homoskedastic, and OLS estimation is efficient. If m_{1n} is heteroskedastic, the usual robust variance estimator can be used. However, even in this case, the degree of heteroskedasticity in u is likely to be small. First, $E(Me) = 0$ as $E(\kappa_{M,e}) = 0$ in $Me = \kappa_{M,e} m_{1,n}$ is a weighted average of e with the same weighting for all observations. As a consequence, OLS is nearly efficient in finite samples, as $\rho^* \kappa_{M,e} m_{1,n} + e \approx e$ in the structural model form in (9) and $\lambda^* \kappa_Z w_{1,n} + e \approx e$ in the reduced model form in (12) even with medium-sized samples. In large samples, these statements will be exactly true and OLS is efficient.

Thus, for estimation of, and inference on, the structural slopes, OLS estimation of the structural model is sufficient. However, if M is known and available to the researcher, an alternative approach is to include m_{1n} as an additional variable in the regression: that is, regress y on Wy, X and m_{1n}. In this case, the regression residual reduces to $\hat{e}$, and the regression is homoskedastic regardless of the nature of m_{1n}. Defining $\tilde{Z} = (Z, m_{1n})$, an estimate of the OLS variance of this regression, $\hat{\zeta} = (\hat{\theta}', \hat{\delta})'$, where δ is the coefficient on m_{1n}, can be obtained as $\hat{V}(\hat{\zeta}) = s_e^2 \tilde{Z}'\tilde{Z}$, where $s_e^2 = \hat{e}'\hat{e}/(N - K - 1)$, and s_e^2 is a consistent estimator for σ_e^2.

3 Panel data

3.1 General considerations

With panel data at hand, the observation index is $j = \{it\}$ with $i = 1, ..., N$ and $t = 1, ..., T$. In general, sorting the data first by t and then by i, we have $V = diag_{t=1,...,T} V_t$ for $V \in \{W, M\}$. In the case of balanced data, $n = NT$, and there are t-invariant $N \times N$ network matrices V_N in each t, so that we have $V = \iota_T \otimes V_N$ for $V \in \{W, M\}$.

We will particularly think of the case where T is large (and not necessarily pertains to time). Subsection 3.2 provides a particular motivation and example for this context. With large T it is reasonable to postulate the disturbances to be

$$e_t = \mu_N + \eta_t \iota_N + \varepsilon_t, \tag{13}$$

where μ_N is an $N \times 1$ vector of i-specific effects which can be treated as fixed or random, depending on the problem at hand, η_t is a t-specific effect which can be treated as fixed or random, and ε_t is an $N \times 1$ vector of IID disturbances.

The model considered for a particular instance of t then reads

$$y_t = \lambda W_{tN} y_t + \alpha \iota_N + X_t \beta + u_t, \quad u_t = \rho M_t u_t + e_t, \tag{14}$$

$$y_t = (I_N - \lambda W_{tN})^{-1} (\alpha \iota_N + X_t \beta + u_t), \quad u_t = (I_N - \rho M_t)^{-1} e_t. \tag{15}$$

We can stack the data, sorting them first by t and then i. For this, we use $\{y, X, u, e, \varepsilon\}$ for stacked vectors and matrices with $n = TN$ rows, considering the balanced case. Specifically,

$$e = \iota_T \otimes \mu_N + \eta_T \otimes \iota_N + \varepsilon, \tag{16}$$

and

$$y = \lambda W y + \alpha \iota_n + X \beta + u, \quad u = \rho M u + e, \tag{17}$$

$$y = (I_n - \lambda W)^{-1} (\alpha \iota_n + X \beta + u), \quad u = (I_n - \rho M)^{-1} e, \tag{18}$$

where $V = diag_{t=1,...,T} \otimes V_{tN}$ for $V \in \{W, M\}$ to consider the case of t-variant network matrices in spite of the panel being balanced.

In the light of the results of Section 2, we can state that

$$\alpha \iota_n + u = \iota_T \otimes \mu_N^\circ + \eta_T^\circ \otimes \iota_N + \varepsilon. \tag{19}$$

Hence, the structural model can be rewritten as

$$y = \lambda W y + X\beta + \iota_T \otimes \mu_N^{\circ} + \eta_T^{\circ} \otimes \iota_N + \varepsilon. \quad (20)$$

And, noting that

$$(I_n + \lambda^* W)(X\beta + \iota_T \otimes \mu_N^{\circ} + \eta_T^{\circ} \otimes \iota_N + \varepsilon) = \iota_T \otimes \mu_N^{\circ\circ} + \eta_T^{\circ\circ} \otimes \iota_N + \varepsilon, \quad (21)$$

the reduced-form model can be rewritten as

$$y = X\beta + \iota_T \otimes \mu_N^{\circ\circ} + \eta_T^{\circ\circ} \otimes \iota_N + \varepsilon. \quad (22)$$

Clearly, the above provides the following insights. First, for X varying in both dimensions i and t, β can be identified by OLS upon treating $\{\mu_N^{\circ}, \mu_N^{\circ\circ}\}$ as well as $\{\eta_N^{\circ}, \eta_N^{\circ\circ}\}$ as fixed. Second, this is the case, irrespective of whether $\{\eta_N^{\circ}, \eta_N^{\circ\circ}\}$ is treated as random or fixed, as long as $E[X(\eta_T \otimes \iota_N)] = 0$. Third, in the latter case, λ can be identified from the structural model as well. Whether treating μ_N (and $\{\mu_N^{\circ}, \mu_N^{\circ\circ}\}$) as fixed or random is immaterial for the identification question, if $E[X(\iota_T \otimes \mu_N)] = 0$, as had been assumed.

3.2 Special case: linearizations of gravity models at a trade elasticity of zero

Behrens et al. (2012) showed that customary gravity models of the class discussed in Arkolakis et al. (2012) – which includes most quantitative trade models used for policy analysis, inter alia the ones of Eaton and Kortum (2002), Anderson and van Wincoop (2003), Melitz (2003), and Bergstrand et al. (2013) – can be log-linearized around a price elasticity of 0 to obtain a structural model form for importer i' regarding its imports from all N exporters as

$$y_{i'N} = \alpha\iota_N + +\lambda W_N y_{i'N} + (\lambda - 1)x_{1N} + \lambda(I_N - W_N)x_{2,i'N} + u_{i'N}, \quad (23)$$

where the original data carry indices $\{i, i'\} = 1, ..., N$ for exporters and importers, and $y_{i'N}$ is an $N \times 1$ vector of normalized bilateral export flows of all regions $i = 1, ..., N$ to one specific importing region i', α is a constant, λ is the prevailing trade elasticity (the price elasticity of demand; a key structural parameter in trade models), $x_{1N} = (x_{1i})$ is an $N \times 1$ vector of known unit production costs in all regions, and $x_{2,i'N} = (x_{2,ii'})$ is an $N \times 1$ vector of ad-valorem bilateral trade trade costs.[1] $u_{i'N}$ is a residual vector which includes a component pertaining to the linearization error of the nonlinear structural gravity-type trade model.

[1] These trade costs can be further parameterized, as a linear function of log distance, log ad-valorem tariff equivalent, and binary indicators for common culture, language, and history, etc., without any loss of generality. Then, we would have a specification, where $\lambda x_{2,i'N} = \sum_{h=1}^{H} \beta_h z_{h,i'N}$, with $z_{h,i'N}$ being one of H-many mentioned observable trade-cost measures and β_h being the corresponding parameter on it.

Egger and Staub (2025) demonstrate that the $N \times N$ network matrix W_N corresponds to the Sherman-Morrison-Woodbury class $W_N = \iota_N w'_{2,N}$ with $w_{2,N}$ being an $N \times 1$ vector of population (or employment shares), when adopting this linearization.

Stacking the data for all exporters i, we arrive at a model for $n = N^2$ cross-sectional bilateral sales data (including domestic sales from i to i'):

$$y = \alpha\iota_n + \lambda W y + (\lambda - 1)x_1 + \lambda(I_n - W)x_2 + u. \tag{24}$$

In light of the discussion of generic panel models, this is a special case, where the network matrix for the stacked (panel) data is constructed based on a Kronecker product, $W = (I_N \otimes W_N)$, as the elements of W_N are invariant across exporter units i. We refer the reader to Egger and Staub (2025) for further details on particular estimation issues with this model. However, what is key is that the network matrix of the model has the special form addressed here, which leads to a reduced model form which is linear in parameters. This had not been recognized in Behrens et al. (2012) which, in our opinion, led to unnecessary complications for estimation and inference, there.

4 Monte Carlo experiments

4.1 Design

For the Monte Carlo simulations, we consider twelve alternative parameter and network designs, which are summarized in Table 2. All designs involve network matrices $V_N = v_{1,N}v'_{2,N}$, where $v'_{2,N}$ is an $1 \times N$ vector of shares with the property $v'_{2,N}\iota_N = 1$. V_N is either row-normalized, when $v_{1,N} = \iota_N$, or it is maximum-row-sum normalized when the elements of $v_{1,N}$ are $v_{1,N,i} = (0, 1]$. We use each of the twelve designs both in a cross-sectional setting, where $n = N$ and $V = V_N$, and also in a panel setting with $n = NT$ and $V = I_T \otimes V_N$. All designs involve a vector of random IID error terms ε with a variance-covariance matrix $E(\varepsilon\varepsilon') = \sigma^2_\varepsilon I_n$ and $\sigma^2_\varepsilon = 1$.

With panel data, we also generate IID variables μ_N and η_T with $E(\mu_N\mu'_N) = \sigma^2_\mu I_N$, $E(\eta_T\eta'_T) = \sigma^2_\eta I_T$ and $\sigma^2_\mu = \sigma^2_\eta = 1$, which are mutually fully independent and also fully independent of ε. Hence, there would be no need to use fixed effects to estimate β—the parameters on X—consistently. We draw the columns of $\{\varepsilon, \mu_N, \eta_T\}$ as $\varepsilon \sim \mathcal{N}(0\iota_n, I_n)$, $\mu_N \sim \mathcal{N}(0\iota_N, I_N)$, and $\eta_T \sim \mathcal{N}(0\iota_T, I_T)$, respectively.

All designs involve two regressors $X = (x_1, x_2)$, where $x_k = 2\iota_n + \xi_k$, which we draw fully independently of all other variables and of each other. Specifically, using $x = vec(X)$, we draw x as fixed in repeated samples with a normal distribution of $x \sim \mathcal{N}(2\iota_{2n}, 4I_{2n})$. That is, we draw x once for our cross-sectional and once for our panel simulations, and keep the same x across all replications and designs. The true vector of parameters, including the constant and slope parameters on X, is

$(\alpha, \beta') = (3, 5, 5)$, which we also keep fixed across designs. All models include a network structure in y, $\bar{y} = Wy$ and feature the key network parameter $\lambda \neq 0$. Half of the twelve designs also include a network structure in the error, $\bar{u} = Wu$, with corresponding parameter $\rho \neq 0$.

In the cross-sectional model, we consider sample sizes of $N = \{200; 400\}$. For the panel model, we use $N = 200$ and let $T = \{50, 100\}$. Hence, with panel data, $n = \{5,000; 10,000\}$. Such a choice of T reflects common values in macro-econometric applications involving countries as cross-sectional units. In the special case of the mentioned linearization of gravity models of trade between countries, panels are quadratic with $T = N$. While quadratic panels are considered further in Egger and Staub (2025), here we explore rectangular macro panels in which $T < N$. However, motivated by the linearized gravity model, we also consider a row-normalized network design of $W_N = \iota_N w'_{2,N}$.

Table 2: Monte Carlo Simulation Designs

		Network structure in u ($\rho \neq 0$)						No network structure in u ($\rho = 0$)					
		D1	D2	D3	D4	D5	D6	D7	D8	D9	D10	D11	D12
Slope parameter	(On variable)												
λ	(Wy)	0.7	2.0	-10	[0.7]	[2.0]	[-10]	0.7	2.0	-10	[0.7]	[2.0]	[-10]
β_1	(x_1)	5.0	5.0	5.0	5.0	5.0	5.0	5.0	5.0	5.0	5.0	5.0	5.0
β_2	(x_2)	5.0	5.0	5.0	5.0	5.0	5.0	5.0	5.0	5.0	5.0	5.0	5.0
Further DGP features													
$w_{1,N}$		$\sim U_N$	$\sim U_N$	$\sim U_N$	ι_N	ι_N	ι_N	$\sim U_N$	$\sim U_N$	$\sim U_N$	ι_N	ι_N	ι_N
ρ		0.5	-4.0	-1.0	0.5	-4.0	-1.0	—	—	—	—	—	—
$m_{1,N}$		ι_N	ι_N	ι_N	$\sim U_N$	$\sim U_N$	$\sim U_N$	—	—	—	—	—	—
α		3.0	3.0	3.0	[3.0]	[3.0]	[3.0]	3.0	3.0	3.0	[3.0]	[3.0]	[3.0]

We report results based on $1,000$ Monte Carlo replications for the key parameters of interest, the structural slope parameters of the model λ, β_1, β_2. In each replication, the model is estimated by OLS as a linear regression of y on Wy, x_1, and x_2. When $w_{1N} = \iota_N$, the network-weighted variable Wy is constant and therefore collinear with the constant term in the regression, α. Hence, as mentioned above, in this case, λ is not identified. For designs with this feature (D4-D6 and D10-D12), we exclude the constant vector ι_N from the regression, letting the parameter on Wy absorb the joint parameters on the constant terms, and present results only for the identified parameters β_1 and β_2, the coefficients on x_1 and x_2.

4.2 Statistics

Let us use $\zeta \in \{\lambda, \beta_1, \beta_2, \alpha\}$ to refer to a generic candidate model parameter of interest. And let $\hat{\zeta}$ pertain to an estimate thereof. Moreover, let us use $\{\hat{\zeta}_{p25}, \hat{\zeta}_{p50}, \hat{\zeta}_{p75}\}$ as the 25th, 50th (median), and 75th percentiles of an estimate across the $1,000$ Monte Carlo draws. Then we can define the robust bias and root mean squared error of an estimate, respectively, as

$$\text{RBias}(\hat{\zeta}) = \hat{\zeta}_{p50} - \zeta, \quad \text{RRMSE}(\hat{\zeta}) = \left[\text{RBias}^2(\hat{\zeta}) + \left(\frac{\hat{\zeta}_{p75} - \hat{\zeta}_{p25}}{1.35}\right)^2\right]^{1/2}. \tag{25}$$

These two measures are known to be robust to occasional outliers in the Monte Carlo vector of point estimates relative to their average-based alternatives.

Regarding inference, we provide two statistics. First, we report on the average standard error for a parameter estimated across the Monte Carlo draws within a design ($\bar{se}$) relative to the standard deviation across the Monte Carlo draws (sd). We expect this ratio to be close to one and converge towards one as the sample size increases with a consistent estimator. Second, we report on the rejection rate using a nominal test size of 0.05 ($RF_{0.05}$), and we expect the corresponding statistics to be close to 0.05 for parameters that are identified and whose standard errors are estimated consistently.

4.3 Results

Cross sections:

Table 3 displays the estimated robust biases and robust root mean squared errors for the cross-sectional data-generating processes, with the top panel containing the results from the cases with $N = 200$ and the bottom panel those for $N = 400$. For $\hat{\beta}_k$, the estimates of the coefficients on the covariates x_k, the biases are very small for all twelve DGPs even in the case of the smaller sample size of 200 observations. The root mean squared errors are therefore essentially driven only by the sampling variance of these estimates. As one would expect when doubling the sample size, the standard deviation and consequently the RRMSE shrink approximately by a factor of $\sqrt{2}$ from the top relative to the bottom panel. For instance, the RRMSE of $\hat{\beta}_1$ in D1 decreases from 0.0350 to 0.0257. Similar results are observed for the network coefficient $\hat{\lambda}$ in the DGPs, where this parameter is separately identified from the model constant α, D1–D3 and D7–D9: the coefficients are estimated virtually without bias, and their RRMSE behaves as expected. This corroborates the conclusions drawn in Section 2.

The results from the table indicate that the strength and direction of the network structure of the error u make very little difference to the quality of the estimates of β_k—there is almost no difference in the results across D1–D3. More broadly, the results show that even whether the model error u exhibits a network structure at all (D1–D6) or not (D7–D8) has negligible effects on the estimates. The same holds for the estimates of λ, where they are available.

Tables 4 and 5 examine the quality of the inference in the data generating processes considered in this simulation. Table 4 shows the results using conventional standard errors, while the statistics in Table 5 are based on heteroskedasticity-robust standard errors. The results in both tables are quite similar. The ratio of average standard error to the standard deviation of the estimator is close to 1 for both $\hat{\beta}_k$. While these statistics are subject to some simulation error, it is remarkable that across DGPs in both tables the ratio is always within 2% of 1. The results for $\hat{\lambda}$ are almost as good, also often being very close to 1 and with the largest deviations from 1 being around 3.5%.

A similar picture emerges from considering the estimated rejection frequencies, $RF_{0.05}$. These are close to the nominal size of 0.05 for all DGPs for the key parameters $\hat{\lambda}$, $\hat{\beta}_1$ and $\hat{\beta}_2$, and are not statistically significantly different from 0.05 at the 5% significance level (95% confidence intervals range from about $RF_{0.05} \pm 0.012$ to $RF_{0.05} \pm 0.015$).

Panels:

The results for the simulation with panel data DGPs are shown in Tables 6–8, following the same structure as the cross-sectional results. While, as mentioned above, in the true DGP the unit-specific and time-specific effects are uncorrelated to the model variables and therefore 'random effects', the estimation treats them as 'fixed effects', that is, as parameters to be estimated, since in practice it is not known to the researcher *ex ante* whether or not these effects are independent or not of the other covariates. This implies that $N + T$ additional degrees of freedom are lost as these fixed effects are accounted for in the estimation. The tables of results again feature two panels related to the total number of observations. The top panel is for $T = 50$ and the botton panel for $T = 100$ (and in both cases, $N = 200$).

The results show that, similar to the cross-sectional case, the estimation of the coefficients of the x_k variables displays virtually no bias and low root mean squared error throughout all DGP designs (Table 6). For the DGPs where λ can be estimated, the panel scenario also shows mostly low bias and RMSE. Interestingly, for DGP D3, there is a somewhat larger RMSE, although it shrinks when T is increased, as expected. The corresponding DGP without network structure in u, D9, also has a higher RMSE than the other two DGPs, D7 and D8.

The patterns observed for the results regarding inference on the estimated parameters in Tables 7 and 8 are largely in line with the results obtained for the estimates in Table 6 as well as the results for inference in the cross-sectional DGPs. For the parameters $\hat{\beta}_k$, both the ratio of standard errors to standard deviation is consistently close to 1 across all DGPs and sample sizes, and the empirical size of the test of the null hypothesis $\beta_k = \beta_k^0$ is close to and not statistically significantly different from the nominal level of 5%. Using default standard errors in Table 7 or cluster-and-heteroskedasticity-robust standard errors Table 8 makes little difference in these DGPs, as expected. We find similar results for inference on the key network parameter λ. The ratio $\bar{se}/sd$ as well as the estimated rejection frequencies are close to their theoretical values. This even includes the DGPs D3 and D7, which showed signs of moderate imprecision in the previous table. The results in the present tables show that this increased variance does not hamper inference for λ.

Overall, the simulations of the cross-sectional and panel data DGPs have shown that the finite sample properties of the estimators for the key variables in the model, namely the network effect

λ as well as the effect of covariates β, displayed a good performance over all simulated designs, including models with an additional network structure in the error. While this error network structure has not been taken into account in estimation, the simulation results demonstrate that this has virtually no effect on the quality of the esitmators' performance.

5 Conclusions

This paper explores the use of network matrices of a Sherwood-Morrison-Woodbury form in econometrics. Such matrices feature self-reflection in that the diagonal elements are generally nonzero. They share this property with another well-known network matrix, namely the Leontief matrix used in national and global input-output analysis. However, Sherwood-Morrison-Woodbury network matrices of size $N \times N$ can be formulated in terms of just two $N \times 1$ vectors. They are generally asymmetric. Other than having nonzero elements, what distinguishes these matrices from the network matrices typically used in econometrics is that their entries are defined directly in the real-number space without normalization, and the same holds for the corresponding peer-effects parameters (often referred to as SAR or SARAR parameters in spatial and network econometrics). However, their inverse exists in relatively general circumstances.

We highlight specific issues with the estimation of structural network-model parameters in models that are linear in their structural form. Moreover, we demonstrate that identifiable parameters of interest can be estimated by OLS in cross-sectional (IID) data and by OLS or GLS in panel data settings in small to moderately-sized data sets. Finally, we provide reasoning for panel data, why Sherman-Morrison-Woodbury inverses would emerge in practice in the context of linearized place-to-place flow models. The focus of the example is on trade but with similar conclusions holding for worker or household mobility equations.

References

Anderson JE, van Wincoop E. 2003. Gravity with gravitas: a solution to the border puzzle. *American Economic Review* **93**: 170–192.

Arkolakis C, Costinot A, Rodríguez-Clare A. 2012. New trade models, same old gains? *American Economic Review* **102**: 94–130.

Bartlett MS. 1951. An inverse matrix adjustment arising in discriminant analysis. *The Annals of Mathematical Statistics* **22**: 107–111.

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

Bergstrand JH, Egger PH, Larch M. 2013. Gravity redux: estimation of gravity-equation coefficients, elasticities of substitution, and general equilibrium comparative statics under asymmetric bilateral trade costs. *Journal of International Economics* **89**: 110–121.

Debarsy N. 2012. The mundlak approach in the spatial durbin panel data model. *Spatial Economic Analysis* **7**: 109–131.

Eaton J, Kortum SS. 2002. Technology, geography, and trade. *Econometrica* **70**: 1741–1779.

Egger PH, Staub KE. 2025. Estimating trade network models without network econometrics. *Working Paper* .

Hao Y, Simoncini V. 2021. The Sherman-Morrison-Woodbury formula for generalized linear matrix equations and applications. *Numerical Linear Algebra with Applications* **28**: e2384.

Kapoor M, Kelejian HH, Prucha IR. 2007. Panel data models with spatially correlated error components. *Journal of Econometrics* **140**: 97–130.

Kelejian H, Piras G. 2017. *Spatial Econometrics*. Academic Press.

Kelejian H, Piras G. 2020. Spillover effects in spatial models: Generalizations and extensions. *Journal of Regional Science* **60**: 425–442.

Kelejian HH. 2008. A spatial j-test for model specification against a single or a set of non-nested alternatives. *Letters in Spatial and Resource Sciences* **1**: 3–11.

Kelejian HH, Piras G. 2011. An extension of kelejian's j-test for non-nested spatial models. *Regional Science and Urban Economics* **41**: 281–292.

Kelejian HH, Piras G. 2014. Estimation of spatial models with endogenous weighting matrices, and an application to a demand model for cigarettes. *Regional Science and Urban Economics* **46**: 140–149.

Kelejian HH, Piras G. 2016. An extension of the j-test to a spatial panel data framework. *Journal of Applied Econometrics* **31**: 387–402.

Kelejian HH, Piras G. 2018. Important overlooked ivs in spatial models. *Empirical Economics* **55**: 69–83.

Kelejian HH, Prucha IR. 1997. Estimation of spatial regression models with autoregressive errors by two-stage least squares procedures: a serious problem. *International regional science review* **20**: 103–111.

Kelejian HH, Prucha IR. 1998. A generalized spatial two-stage least squares procedure for estimating a spatial autoregressive model with autoregressive disturbances. *Journal of Real Estate Finance and Economics* **17**: 99–121.

Kelejian HH, Prucha IR. 1999. A generalized moments estimator for the autoregressive parameter in a spatial model. *International Economic Review* **40**: 509–533.

Kelejian HH, Prucha IR. 2001. On the asymptotic distribution of the moran i test statistic with applications. *Journal of Econometrics* **104**: 219–257.

Kelejian HH, Prucha IR. 2004. Estimation of simultaneous systems of spatially interrelated cross sectional equations. *Journal of Econometrics* **118**: 27–50.

Kelejian HH, Prucha IR. 2007. Hac estimation in a spatial framework. *Journal of Econometrics* **140**: 131–154.

Kelejian HH, Prucha IR. 2010. Specification and estimation of spatial autoregressive models with autoregressive and heteroskedastic disturbances. *Journal of Econometrics* **157**: 53–67.

Kelejian HH, Prucha IR, Yuzefovich Y. 2004. Instrumental variable estimation of a spatial autoregressive model with autoregressive disturbances: large and small sample result. In LeSage J, Pace RK (eds.) *Advances in Econometrics*. New York: Elsevier, 163–198.

Kelejian HH, Robinson DP. 1992. Spatial autocorrelation: A new computationally simple test with an application to per capita county police expenditures. *Regional Science and Urban Economics* **22**: 317–331.

Kelejian HH, Robinson DP. 1998. A suggested test for spatial autocorrelation and/or heteroskedasticity and corresponding monte carlo results. *Regional Science and Urban Economics* **28**: 389–417.

Kelejian HH, Robinson DP. 2004. The influence of spatially correlated heteroskedasticity on tests for spatial correlation. In *Advances in Spatial Econometrics: Methodology, Tools and Applications*. Springer, 79–97.

Melitz MJ. 2003. The impact of trade on intra-industry reallocations and aggregate industry productivity. *econometrica* **71**: 1695–1725.

Riedel KS. 1992. A Sherman–Morrison–Woodbury identity for rank augmenting matrices with application to centering. *SIAM Journal on Matrix Analysis and Applications* **13**: 659–662.

Sherman J, Morrison WJ. 1950. Adjustment of an inverse matrix corresponding to a change in one element of a given matrix. *The Annals of Mathematical Statistics* **21**: 124–127.

Woodbury MA. 1949. The stability of out-input matrices. *Chicago, IL* .

Woodbury MA. 1950. Inverting modified matrices. *Memorandum Rept. 42, Statistical Research Group, Princeton University, Princeton, NJ* .

Tables

Table 3: Simulation results: OLS estimates of cross-sectional network model.

		Network structure in u ($\rho \neq 0$)						No network structure in u ($\rho = 0$)					
		D1	D2	D3	D4	D5	D6	D7	D8	D9	D10	D11	D12
I. $N = 200$													
$\hat{\lambda}$	RBias	-0.0002	0.0000	-0.0022	—	—	—	-0.0002	0.0000	-0.0022	—	—	—
	RRMSE	0.0069	0.0000	0.0640	—	—	—	0.0069	0.0000	0.0638	—	—	—
$\hat{\beta}_1$	RBias	-0.0001	-0.0001	-0.0001	-0.0001	-0.0000	0.0000	-0.0001	-0.0001	-0.0001	-0.0001	-0.0001	-0.0001
	RRMSE	0.0350	0.0350	0.0350	0.0350	0.0343	0.0351	0.0350	0.0350	0.0350	0.0349	0.0349	0.0349
$\hat{\beta}_2$	RBias	0.0024	0.0024	0.0024	0.0029	0.0029	0.0029	0.0024	0.0024	0.0024	0.0030	0.0030	0.0030
	RRMSE	0.0378	0.0378	0.0378	0.0376	0.0375	0.0374	0.0378	0.0378	0.0378	0.0375	0.0375	0.0375
II. $N = 400$													
$\hat{\lambda}$	RBias	-0.0000	-0.0000	-0.0003	—	—	—	-0.0000	-0.0000	-0.0003	—	—	—
	RRMSE	0.0051	0.0005	0.0431	—	—	—	0.0051	0.0005	0.0432	—	—	—
$\hat{\beta}_1$	RBias	-0.0008	-0.0008	-0.0008	-0.0008	-0.0007	-0.0009	-0.0008	-0.0008	-0.0008	-0.0008	-0.0008	-0.0008
	RRMSE	0.0257	0.0257	0.0257	0.0251	0.0254	0.0254	0.0257	0.0257	0.0257	0.0249	0.0249	0.0249
$\hat{\beta}_2$	RBias	0.0003	0.0003	0.0003	0.0005	0.0005	0.0005	0.0003	0.0003	0.0003	0.0005	0.0005	0.0005
	RRMSE	0.0261	0.0261	0.0261	0.0259	0.0259	0.0259	0.0261	0.0261	0.0261	0.0259	0.0259	0.0259

Notes: Results obtained by Monte Carlo simulation over 1000 replications for DGP designs D1-D2 (see Table 2). Further details on the data generating process can be found in Section 4.3. RBias and RRMSE refer to simulation estimates of the robust bias and robust root mean squred error as defined in eq. (25).

Table 4: Simulation results: Inference for cross-sectional network model.

		Network structure in u ($\rho \neq 0$)						No network structure in u ($\rho = 0$)					
		D1	D2	D3	D4	D5	D6	D7	D8	D9	D10	D11	D12
I. $N = 200$													
$\hat{\lambda}$	$\bar{se}/sd$	1.008	1.008	1.009	—	—	—	1.009	1.008	1.009	—	—	—
	$RF_{0.05}$	0.041	0.041	0.041	—	—	—	0.041	0.040	0.041	—	—	—
$\hat{\beta}_1$	$\bar{se}/sd$	0.992	0.992	0.992	0.993	0.995	0.995	0.992	0.992	0.992	0.994	0.994	0.994
	$RF_{0.05}$	0.048	0.048	0.048	0.048	0.047	0.049	0.048	0.048	0.048	0.048	0.048	0.048
$\hat{\beta}_2$	$\bar{se}/sd$	1.011	1.011	1.011	1.012	1.013	1.012	1.011	1.011	1.011	1.012	1.012	1.012
	$RF_{0.05}$	0.045	0.045	0.045	0.044	0.045	0.044	0.045	0.045	0.045	0.044	0.044	0.044
II. $N = 400$													
$\hat{\lambda}$	$\bar{se}/sd$	0.966	0.966	0.966	—	—	—	0.966	0.966	0.966	—	—	—
	$RF_{0.05}$	0.061	0.061	0.061	—	—	—	0.061	0.061	0.061	—	—	—
$\hat{\beta}_1$	$\bar{se}/sd$	1.003	1.003	1.003	1.007	1.003	1.005	1.003	1.003	1.003	1.006	1.006	1.006
	$RF_{0.05}$	0.055	0.055	0.055	0.053	0.054	0.055	0.055	0.055	0.055	0.055	0.055	0.055
$\hat{\beta}_2$	$\bar{se}/sd$	0.982	0.982	0.982	0.982	0.983	0.982	0.982	0.982	0.982	0.982	0.982	0.982
	$RF_{0.05}$	0.051	0.051	0.051	0.050	0.050	0.050	0.051	0.051	0.051	0.050	0.050	0.050

Notes: Results obtained by Monte Carlo simulation over 1000 replications for DGP designs D1-D2 (see Table 2). Further details on the data generating process can be found in Section 4.3. $\bar{se}/sd$ refers to simulation estimates of the average standard error to standard deviation ratio. $RF_{0.05}$ refers to the estimated rejection frequency of hypothesis tests at the nominal size of 0.05 for $H_0: \theta = 0$, $\theta \in \{\lambda, \beta_1, \beta_2\}$.

Table 5: Simulation results: Heteroskedasticity-robust inference for cross-sectional model.

		Network structure in u ($\rho \neq 0$)						No network structure in u ($\rho = 0$)					
		D1	D2	D3	D4	D5	D6	D7	D8	D9	D10	D11	D12
I. $N = 200$													
$\hat{\lambda}$	$\bar{se}/sd$	1.005	1.005	1.005	—	—	—	1.005	1.005	1.005	—	—	—
	$RF_{0.05}$	0.044	0.043	0.044	—	—	—	0.044	0.043	0.044	—	—	—
$\hat{\beta}_1$	$\bar{se}/sd$	0.986	0.986	0.986	0.987	0.989	0.989	0.986	0.986	0.986	0.988	0.988	0.988
	$RF_{0.05}$	0.055	0.055	0.055	0.052	0.052	0.049	0.055	0.055	0.055	0.050	0.050	0.050
$\hat{\beta}_2$	$\bar{se}/sd$	1.005	1.005	1.005	1.006	1.006	1.006	1.005	1.005	1.005	1.006	1.006	1.006
	$RF_{0.05}$	0.043	0.043	0.043	0.042	0.044	0.044	0.043	0.043	0.043	0.043	0.043	0.043
II. $N = 400$													
$\hat{\lambda}$	$\bar{se}/sd$	0.964	0.964	0.964	—	—	—	0.964	0.964	0.964	—	—	—
	$RF_{0.05}$	0.061	0.061	0.061	—	—	—	0.061	0.061	0.061	—	—	—
$\hat{\beta}_1$	$\bar{se}/sd$	0.997	0.997	0.997	1.000	0.997	0.999	0.997	0.997	0.997	1.000	1.000	1.000
	$RF_{0.05}$	0.061	0.061	0.061	0.060	0.060	0.060	0.061	0.061	0.061	0.060	0.060	0.060
$\hat{\beta}_2$	$\bar{se}/sd$	0.978	0.978	0.978	0.978	0.979	0.978	0.978	0.978	0.978	0.978	0.978	0.978
	$RF_{0.05}$	0.056	0.056	0.056	0.056	0.055	0.055	0.056	0.056	0.056	0.056	0.056	0.056

Notes: Results obtained by Monte Carlo simulation over 1000 replications for DGP designs D1-D2 (see Table 2). Further details on the data generating process can be found in Section 4.3. $\bar{se}/sd$ refers to simulation estimates of the average standard error to standard deviation ratio. $RF_{0.05}$ refers to the estimated rejection frequency of hypothesis tests at the nominal size of 0.05 for $H_0: \ \theta = 0$, $\theta \in \{\lambda, \beta_1, \beta_2\}$.

Table 6: SIMULATION RESULTS: OLS ESTIMATES OF PANEL NETWORK MODEL ($N = 200$).

		Network structure in u ($\rho \neq 0$)						No network structure in u ($\rho = 0$)					
		D1	D2	D3	D4	D5	D6	D7	D8	D9	D10	D11	D12
I. $T = 50$													
$\hat{\lambda}$	RBias	0.0004	0.0001	-0.0047	—	—	—	-0.0000	0.0000	-0.0002	—	—	—
	RRMSE	0.0106	0.0011	0.1841	—	—	—	0.0155	0.0009	0.1506	—	—	—
$\hat{\beta}_1$	RBias	0.0003	0.0002	0.0002	-0.0004	0.0015	0.0009	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003
	RRMSE	0.0047	0.0047	0.0047	0.0047	0.0049	0.0047	0.0047	0.0047	0.0047	0.0047	0.0047	0.0047
$\hat{\beta}_2$	RBias	0.0001	0.0001	0.0001	-0.0007	0.0015	0.0009	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001
	RRMSE	0.0049	0.0050	0.0050	0.0051	0.0052	0.0050	0.0050	0.0050	0.0050	0.0050	0.0050	0.0050
II. $T = 100$													
$\hat{\lambda}$	RBias	0.0001	0.0000	0.0024	—	—	—	0.0007	0.0001	0.0056	—	—	—
	RRMSE	0.0075	0.0017	0.1148	—	—	—	0.0115	0.0014	0.0953	—	—	—
$\hat{\beta}_1$	RBias	-0.0002	-0.0002	-0.0002	0.0004	-0.0013	-0.0008	-0.0002	-0.0002	-0.0002	-0.0002	-0.0002	-0.0002
	RRMSE	0.0035	0.0035	0.0035	0.0035	0.0037	0.0036	0.0035	0.0035	0.0035	0.0035	0.0035	0.0035
$\hat{\beta}_2$	RBias	0.0000	0.0000	0.0000	-0.0015	0.0029	0.0015	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
	RRMSE	0.0035	0.0035	0.0035	0.0038	0.0046	0.0038	0.0035	0.0035	0.0035	0.0035	0.0035	0.0035

Notes: Results obtained by Monte Carlo simulation over 1000 replications for DGP designs D1-D2 (see Table 2). Further details on the data generating process can be found in Section 4.3. RBias and RRMSE refer to simulation estimates of the robust bias and robust root mean squred error as defined in eq. (25).

Table 7: SIMULATION RESULTS: INFERENCE FOR PANEL NETWORK MODEL ($N = 200$).

		Network structure in u ($\rho \neq 0$)						No network structure in u ($\rho = 0$)					
		D1	D2	D3	D4	D5	D6	D7	D8	D9	D10	D11	D12
I. $T = 50$													
$\hat{\lambda}$	$\bar{se}/sd$	1.047	0.989	1.004	—	—	—	1.027	1.027	1.027	—	—	—
	$RF_{0.05}$	0.045	0.050	0.051	—	—	—	0.052	0.052	0.052	—	—	—
$\hat{\beta}_1$	$\bar{se}/sd$	1.032	1.033	1.032	1.045	1.077	1.044	1.032	1.032	1.032	1.032	1.032	1.032
	$RF_{0.05}$	0.043	0.044	0.044	0.043	0.044	0.042	0.043	0.043	0.043	0.043	0.043	0.043
$\hat{\beta}_2$	$\bar{se}/sd$	1.023	1.022	1.022	1.035	1.065	1.033	1.022	1.022	1.022	1.022	1.022	1.022
	$RF_{0.05}$	0.046	0.047	0.047	0.046	0.043	0.042	0.046	0.046	0.046	0.047	0.047	0.047
II. $T = 100$													
$\hat{\lambda}$	$\bar{se}/sd$	0.996	0.948	0.954	—	—	—	0.971	0.971	0.971	—	—	—
	$RF_{0.05}$	0.048	0.061	0.067	—	—	—	0.058	0.058	0.058	—	—	—
$\hat{\beta}_1$	$\bar{se}/sd$	1.015	1.015	1.015	1.036	1.084	1.033	1.015	1.015	1.015	1.015	1.015	1.015
	$RF_{0.05}$	0.044	0.044	0.044	0.045	0.052	0.046	0.043	0.043	0.043	0.044	0.044	0.044
$\hat{\beta}_2$	$\bar{se}/sd$	0.963	0.963	0.963	0.982	1.029	0.981	0.963	0.963	0.963	0.963	0.963	0.963
	$RF_{0.05}$	0.055	0.057	0.057	0.079	0.106	0.066	0.055	0.055	0.055	0.056	0.056	0.056

Notes: Results obtained by Monte Carlo simulation over 1000 replications for DGP designs D1-D2 (see Table 2). Further details on the data generating process can be found in Section 4.3. $\bar{se}/sd$ refers to simulation estimates of the average standard error to standard deviation ratio. $RF_{0.05}$ refers to the estimated rejection frequency of hypothesis tests at the nominal size of 0.05 for $H_0 : \theta = 0$, $\theta \in \{\lambda, \beta_1, \beta_2\}$.

Table 8: Simulation results: Heteroskedasticity- and cluster-robust inference for panel network model ($N = 200$).

		Network structure in u ($\rho \neq 0$)						No network structure in u ($\rho = 0$)					
		D1	D2	D3	D4	D5	D6	D7	D8	D9	D10	D11	D12
I. $T = 50$													
$\hat{\lambda}$	$\bar{se}/sd$	1.045	0.989	1.004	—	—	—	1.026	1.026	1.026	—	—	—
	$RF_{0.05}$	0.052	0.054	0.058	—	—	—	0.056	0.056	0.056	—	—	—
$\hat{\beta}_1$	$\bar{se}/sd$	1.030	1.030	1.030	1.044	1.084	1.044	1.030	1.030	1.030	1.030	1.030	1.030
	$RF_{0.05}$	0.041	0.041	0.041	0.040	0.041	0.040	0.041	0.041	0.041	0.041	0.041	0.041
$\hat{\beta}_2$	$\bar{se}/sd$	1.020	1.020	1.020	1.034	1.067	1.032	1.020	1.020	1.020	1.020	1.020	1.020
	$RF_{0.05}$	0.042	0.044	0.043	0.051	0.042	0.043	0.042	0.042	0.042	0.044	0.044	0.044
II. $T = 100$													
$\hat{\lambda}$	$\bar{se}/sd$	0.993	0.947	0.952	—	—	—	0.969	0.969	0.969	—	—	—
	$RF_{0.05}$	0.053	0.067	0.062	—	—	—	0.057	0.057	0.057	—	—	—
$\hat{\beta}_1$	$\bar{se}/sd$	1.017	1.016	1.016	1.043	1.103	1.040	1.017	1.017	1.017	1.017	1.017	1.017
	$RF_{0.05}$	0.049	0.050	0.050	0.044	0.047	0.050	0.050	0.050	0.050	0.049	0.049	0.049
$\hat{\beta}_2$	$\bar{se}/sd$	0.963	0.963	0.963	0.981	1.026	0.980	0.963	0.963	0.963	0.963	0.963	0.963
	$RF_{0.05}$	0.058	0.059	0.059	0.083	0.117	0.067	0.059	0.059	0.059	0.058	0.058	0.058

Notes: Results obtained by Monte Carlo simulation over 1000 replications for DGP designs D1-D2 (see Table 2). Further details on the data generating process can be found in Section 4.3. $\bar{se}/sd$ refers to simulation estimates of the average standard error to standard deviation ratio. $RF_{0.05}$ refers to the estimated rejection frequency of hypothesis tests at the nominal size of 0.05 for $H_0: \theta = 0$, $\theta \in \{\lambda, \beta_1, \beta_2\}$.