CEIS Tor Vergata Research Papers

Università di Roma Tor Vergata

ISSN 2610-931X

CEIS Tor Vergata

RESEARCH PAPER SERIES

Vol. 16, Issue 6, No. 439 – October 2018

Estimating Models with Dynamic Network Interactions and Unobserved Heterogeneity

Luisa Corrado and Salvatore Di Novo

Estimating models with dynamic network interactions and unobserved heterogeneity*

Luisa Corrado† and Salvatore Di Novo‡

This Version: July 2018

Abstract

In this paper, we propose an approach to estimate models with network interactions in the presence of individual unobserved heterogeneity. The latter may impact the formation of ties and/or exogenous effects, thereby undermining identification of the associated parameters. In a panel setting, we devise a way to cope with these sources of endogeneity by relying on observable variations. When exogenous effects are involved, one can control for unobserved heterogeneity by including time-averages of the endogenous variables. When unobserved individual traits affect the process of network formation, it is possible to explore the role of network statistics. We derive a 2SLS estimator in order to address simultaneity bias, relying on sources of variation provided by the product between successive powers of the network matrix and the matrix of exogenous covariates; we assess the performances of the method via a Monte Carlo exercise, considering various combination of models and different ranges of parameters for both network interactions and the social multiplier. We also separately assess the cases in which unobserved sources hit the network structure only or act on exogenous effects as well. Focusing on the former case, our approach may be also applied when a simple cross-section is available. More generally, it does not require full knowledge of the spectrum of agents' interactions.

Key-Words: Networks, Individual Unobserved Heterogeneity, Dynamic Network Formation, network Statistics.

JEL: C31, C36.

*We do thank Federico Belotti, Donato Ceci, Alberto Holly, Luis R. Izquierdo, Andrea Piano Mortari, Ugo Zannini for suggestions and comments, as well as participants to the 49th Scientific Meeting of the Italian Statistical Society at University of Palermo.

†University of Rome Tor Vergata, Department of Economics and Finance Department, Italy. E-mail: luisa.corrado@uniroma2.it.

‡University of Rome Tor Vergata, Department of Economics and Finance, Italy. E-mail: salvatore.di.novo@uniroma2.it.

1 Introduction

The estimation of models accounting for interactions across agents has become widely popular in recent years, because of a renewed interest to evaluate the role of spillovers. Those spillovers may account for spatial, social or economic interactions. Indeed, this distinction mainly refers to the underlying phenomenon of interest, not to the statistical method one is willing to use to make inference, though the nature of the phenomenon we are studying and the associated modelling may affect the soundness of estimation results. Whether one is focusing on spatial or network interactions, it is well known that the presence of a term accounting for interactions is not identified by simple least squares estimation because of what is commonly known as "simultaneity bias". In these models, interactions enter the estimating equation as the product between the so-called adjacency (interaction, proximity, etc) matrix and the dependent variable (and/or some L-dimensional vector of covariates), a description which may remind of linear simultaneous equation models. Among others Paula et al. [2018] stress how network formation models present difficulties for identification in this setting, especially when links can be interdependent. Various solutions have been proposed in the literature to tackle simultaneity bias. For instance, Kelejian and Prucha [1998] introduced a so-called generalized spatial 2SLS (GS2SLS) in a cross-section framework, an approach involving successive powers of the proximity matrix applied to those variables producing spillovers. This approach has been subsequently extended by Lee [2003] - to set out optimal instruments - and then extended by Baltagi et al. [2014] in the context of models with multilevel error structures. However, all those studies presume the only source of endogeneity in estimating regression models including interactions variables is the one aforementioned, maintaining that the interaction matrix is being exogenously given.

Studies introducing endogenous interaction matrices started to emerge in the recent years, from Lee et al. [2012] - who focus on the identification of central players in a network, Goldsmith-Pinkham and Imbens [2013] - who study how homophily affect the formation of ties - to Horrace et al. [2016] - who consider the role of selection bias onto the network structure.

In this paper we consider threats posed by unobserved individual heterogeneity when estimating models where endogenous effects (say, interactions) and exogenous effects are present and we do focus on a panel setting. Unobserved heterogeneity is thought to affect exogenous variables on one hand, while it does also impact links forming the interaction matrix, thereby violating any assumption about an exogenously given interaction matrix. As a solution, we propose using observed quantities to control for such heterogeneity. Since the latter hits two different observed "parts" of the estimating model, we i) exploit the panel setting to back-out the time-invariant component of the otherwise exogenoues variables and ii) consider network statistics which are tightly linked with heterogenous profiles. We then use those quantities to proxy unobserved heterogeneity and then proceed using the aforementioned approaches to estimate the parameters of our reference model. Indeed, in i) we are just applying to a spatial/network setting the argument raised by Mundlak [1978] in showing equivalence between the customarily termed "fixed-effects" and "random-effects" estimators; ii) is relatively novel, though Graham

[2017] uses indegree meausures to control for unobserved heterogeneity in the context of network estimation; still, Liu [2014] focuses on Bonacich centrality measures to improve identification of social interactions. More recently, Paula et al.'s [2018] present a framework to identify preference parameters based on sets of local network structures which come from the structure of individual preferences themselves. This work shares underlying similarities with Qu and Lee [2015], who also rely on observed quantities to in order to account for the endogenous formation of network ties. We investigate the finite sample performances of the proposed estimator over various simulation exercises.

The paper is organised as follows: section two presents the general model framework; section three addresses the impact of unobserved heterogeneity on the formation of ties and/or on the exogenous effects and describes our proposed estimator; section four describes the network formation process in presence of unobserved heterogeneity; section five assess the performances of the proposed method via a Monte Carlo exercise considering various combination of models and different ranges of parameters for both network interactions and the social multiplier; section six concludes.

2 Framework

We consider a setup in which network interactions alongside a vector of other covariates (exogenous or not) do explain an outcome variable. Further, we consider the error component having an additive compound structure, that is, it does consist of a time-invariant and individual specific term alongside a time-varying term. Suppose we use a linear regression model in order to estimate the parameters of variables so far, that is:

$$\boldsymbol{y} = \boldsymbol{G}\boldsymbol{y}\rho + \boldsymbol{X}\boldsymbol{\beta} + \boldsymbol{u}\,, \tag{1}$$

where the sample at hand consists of panel data and $\boldsymbol{y}$ is an $NT \times 1$ vector, $\boldsymbol{X}$ is an $NT \times k$ matrix and $\boldsymbol{G}$ is an $NT \times NT$ block diagonal matrix where the diagonal blocks consists of entries considered at a certain time period. $\boldsymbol{u}$ is an $NT \times 1$ vector accounting for all is unobserved to the econometrician, which we can break down as $\boldsymbol{u} = \boldsymbol{\alpha} \otimes \iota_T + \boldsymbol{\varepsilon}$, where $\boldsymbol{\alpha}$ is an $N \times 1$ vector, $\boldsymbol{i}_T$ is a $T \times 1$ vector of ones and $\boldsymbol{\varepsilon}$ is $NT \times 1$ vector.

Cliff and Ord [1973] introduced a model like that in (1), which is customarily termed SAR (spatial autoregressive), irrespectively of the presence of a set of covariates alongside (spatial) lags of the dependent variable.[1,2]

Following Manski [1993, 1995], we define the observed terms in the right-hand side of eq. (1) as

[1] You may also see, among the others, Anselin [1988]. Recent reviews of the literature with discussions on still-standing issues include Corrado and Fingleton[2012] and Gibbons et al. [2015].

[2] Insofar, "lags" refers to the number of successive powers of the adjacency matrix $\boldsymbol{G}$ applied to $\boldsymbol{x}$ entering the instruments matrix.

endogenoues effects ($\boldsymbol{Gy}$) and contextual effects ($\boldsymbol{X}$).[3] However, estimation of such an equation as it stands is going to yield inconsistent parameters estimates because of what is known as "simultaneity" bias, since $\boldsymbol{y}$ appears on both sides of the equation.

Indeed, let us simply rewrite eq. (1) and premultiply it by the matrix $\boldsymbol{G}$ to get[4]:

$$\begin{aligned}
\boldsymbol{Gy} &= \boldsymbol{G}\left[\boldsymbol{I}-\boldsymbol{G}\rho\right]^{-1}(\boldsymbol{X\beta}+\boldsymbol{u}) = \\
&= \sum_{l=1}^{+\infty}\boldsymbol{G^l X\beta}\rho^{l-1} + \sum_{l=1}^{+\infty}\boldsymbol{G^l\alpha}\rho^{l-1} + \sum_{l=1}^{+\infty}\boldsymbol{G^l\varepsilon}\rho^{l-1} = \\
&= \sum_{l=1}^{p}\left(\boldsymbol{G^l X\beta}\rho^{l-1}+\boldsymbol{G^l\alpha}\rho^{l-1}\right) + \sum_{l=p+1}^{+\infty}\left(\boldsymbol{G^l X\beta}\rho^{l-1}+\boldsymbol{G^l\alpha}\rho^{l-1}\right) + \sum_{l=1}^{+\infty}\boldsymbol{G^l\varepsilon}\rho^{l-1} \overset{def}{=} \qquad (2)\\
&\overset{def}{=} \boldsymbol{S_{GX}\phi} + \boldsymbol{S_{G\alpha}\rho} + \boldsymbol{R}\left(\boldsymbol{S_{GX}^{(p+1,\infty)}\phi^{(p+1,\infty)}}\right) + \boldsymbol{R}\left(\boldsymbol{S_{G\alpha}^{(p+1,\infty)}\rho^{(p+1,\infty)}}\right) + \boldsymbol{T}\left(\boldsymbol{S_{G\varepsilon}^{(1,\infty)}\rho^{(1,\infty)}}\right) \overset{def}{=} \\
&\overset{def}{=} \boldsymbol{S_{GX}\phi} + \boldsymbol{S_{G\alpha}\rho} + \boldsymbol{R_{GX}} + \boldsymbol{R_{G\alpha}} + \boldsymbol{T_{G\varepsilon}} \qquad (3)
\end{aligned}$$

where $\boldsymbol{S_{GX}}$, $\boldsymbol{S_{G\alpha}}$ and $\boldsymbol{\phi}$, $\boldsymbol{\rho}$ are, respectively, $NT \times p$ matrices and $p \times 1$ vectors whose products yields, respectively, the two terms entering the first sum in (2). Further: $\boldsymbol{R_{GX}} = \sum_{l=p+1}^{+\infty}\boldsymbol{G^l X\beta}\rho^{l-1}$, $\boldsymbol{R_{G\alpha}} = \sum_{l=p+1}^{+\infty}\boldsymbol{G^l\alpha}\rho^{l-1}$ and $\boldsymbol{T_{G\varepsilon}} = \sum_{l=p+1}^{+\infty}\boldsymbol{G^l\varepsilon}\rho^{l-1}$.

As it stands eq. (3) provides us with a potentially fruitful set of "internally" observed quantities one may use to explain exogenous variation in $\boldsymbol{Gy}$. By internally, we mean that those quantities are obtained from eq (1). The underlying logic of the approach consists in exploiting variation due to the linkages of one's ownBlume, L. and Brock, W. and Durlauf, S. and Ioannides, Y. [2011] linkages.[5] On this ground, we follow Kelejian and Prucha [1998], Baltagi et al. [2014] and consider a 2SLS estimation procedure[6] in order to recover the parameters entering eq (1). To this end, let us first define $\boldsymbol{H = (S_{GX},\, X)}$and let us rewrite together the two "stages" we have described so far as a matter of clarity:

$$\boldsymbol{y} = \boldsymbol{Gy}\rho + \boldsymbol{X\beta} + \boldsymbol{\alpha} + \boldsymbol{\varepsilon} \qquad (4)$$

$$\boldsymbol{Gy} = \boldsymbol{H}\left[\boldsymbol{\phi},\ \boldsymbol{\beta}\right]' + \boldsymbol{S_{G\alpha}\rho} + \boldsymbol{\xi} \qquad (5)$$

where $\boldsymbol{\xi} = \boldsymbol{R_{GX}} + \boldsymbol{R_{G\alpha}} + \boldsymbol{T_{G\varepsilon}}$. Let us now state the following assumptions:

[3] It is also customary to label as contextual effect any term involving pre-multiplication of $\boldsymbol{X}$ with the network matrix $\boldsymbol{G}$ (or any power of it).

[4] Recall that the matrix $[\boldsymbol{I}-\boldsymbol{G}\rho]$ is invertible if $|\rho| < 1$.

[5] In this respect, the description above resembles the so-called "Hausman instruments", see Hausman et al. [1994].

[6] More generally, one could consider any instrumental variable estimator.

$$\mathbb{E}\left[\boldsymbol{\alpha}|\,\boldsymbol{X}\right]=0 \tag{6}$$

$$\mathbb{E}\left[\boldsymbol{\varepsilon}|\,\boldsymbol{X}\right]=0 \tag{7}$$

$$\mathbb{E}\left[\boldsymbol{\varepsilon}|\,\boldsymbol{S_{GX}}\right]=0 \tag{8}$$

$$\mathbb{E}\left[\boldsymbol{\alpha}|\,\boldsymbol{S_{GX}}\right]=0 \tag{9}$$

$$\mathbb{E}\left[\boldsymbol{\xi}|\,\boldsymbol{S_{GX}}\right]=0 \tag{10}$$

$$\mathbb{E}\left[\boldsymbol{Gy}|\,\boldsymbol{S_{GX}}\right]\neq 0 \tag{11}$$

Assumptions (6) and (7) allow to identify the regression coefficients for the set of contextual effects, while the remaining assumptions are needed to consistently estimate the parameter associated with the endogenoues effect. In particular, assumptions (8), (9) and (10) imply mean independence between $\boldsymbol{S_{GX}}$ and any unobserved term entering either equations (4) and (5), while (11) implies the set of instruments for $\boldsymbol{Gy}$ induces exogenous variation on it. The latter assumption should indeed be complemented with one regarding the column rank of the instruments matrix, a point we will return to later. Lastly, note that assumption (6) implies $\mathbb{E}\left[\boldsymbol{S_{GX}}|\,\boldsymbol{S_{G\alpha}}\right]=0$, while (10) implies that $\mathbb{E}\left[\boldsymbol{S_{GX}}|\,\boldsymbol{T_{GX}}\right]=0$ as well.

Provided assumptions (6)- (11) are satisfied, we consider the following Generalized 2SLS estimator for the parameter vector $[\rho,\,\boldsymbol{\beta}]$ in equation (1):

$$\left[\hat{\rho},\,\hat{\boldsymbol{\beta}}\right]'=\left(\boldsymbol{Z}'\boldsymbol{P}\boldsymbol{Z}\right)^{-1}\left(\boldsymbol{Z}'\boldsymbol{P}\boldsymbol{y}\right) \tag{12}$$

where $\boldsymbol{Z}=(\boldsymbol{Gy},\,\boldsymbol{X})$ and $\boldsymbol{P}=\boldsymbol{H}\left(\boldsymbol{H}'\boldsymbol{H}\right)^{-1}\boldsymbol{H}'$.

3 The role of unobserved heterogeneity

Now, we wish to make a step forward in the model we have briefly described so far. We introduce the case in which unobserved and time-invariant components - α in our setting - do affect both the network formation process - hence, any variable embedding network interactions - and some subset of the "other" covariates. Both cases are likely to invalidate both eq (4) and (5). In this respect, we figure out to approach these issues by relying on variation coming from observable quantities which correlate with unobserved heterogeneity. It is important to separate out each of the two effects heterogeneity has on invalidating the aforementioned estimation strategy. The impact on exogenoues effects is a standard omitted variable problem one would face when trying to estimate any regression equation and this obviously has nothing to do with the presence of endogenous effects in the estimating equation. On the other hand, if unobserved heterogeneity does play a role in the process of link formation - by affecting each row $\boldsymbol{G_{i\cdot}}$ of the interaction matrix $\boldsymbol{G}$ - any variable accounting for interaction effects obtained premultiplying the variable itself by the matrix $\boldsymbol{G}$ is going to be affected as well. This point has no relation with the simultaneity issue considered before, for the vector $\boldsymbol{y}$ appearing on both side of equation (1) would harm identification of parameters irrespectively of $\boldsymbol{G}$ being endogenous or not. Motivated by this discussion, we are going to face each of the two sources of endogeneity described earlier separately in what follows.

3.1 Unobserved heterogeneity and exogenous effects

Let us consider the $NT \times k_1$ matrix $[\boldsymbol{x_1},\, \boldsymbol{x_2},\, \ldots,\, \boldsymbol{x_{K_1}}]$, $k_1 \leq k$, or expanding across N and T, $x^{'}_{kit} = [x_{k11}, \ldots,\, x_{kN1},\, \ldots,\, x_{kNT}]^{'}$ and let us for the moment allow $k_1 = 1$, such that we can write the following:

$$x_{it} = f\left(\alpha_i,\, v^x_{it}\right)\,,\, \forall i,\, t \tag{13}$$

where v^x_{it} accounts for all factors affecting x_{it} but $\alpha_i$and let us assume $v^x_{it} \perp \alpha_i$ so we can invoke separability of the function above.

As it stands, eq (13) implies that assumptions (6) and (7) above are violated, the implication being that identification of contextual effects can not be obtained. Indeed, this is a well known issue in applied econometrics and many solutions have been proposed in this respect. Actually, we use the information embedded in the equation above to express the unobserved α in terms of observed quantities. In this respect, let us note that:

$$\begin{aligned}
\frac{\partial x_{it}}{\partial \alpha_i} &= \frac{\partial x_{it+k}}{\partial \alpha_i}\,, \forall i, t, k \\
\frac{\partial x_{it}}{\partial \alpha_i} &\neq \frac{\partial x_{jt}}{\partial \alpha_j}\,, \forall i, j, t
\end{aligned}$$

So, the variation induced in x_{it} is entirely due variation across observations rather than through time, which means we can recast the relationship between x_{it} and α_i in terms of averages of the former, $\bar{x}_i$, that is, we can consider $\mathbb{E}\left(\bar{x}_i | \, \alpha_i\right)$ in place of eq (13). Since the latter implies some dependence relation between $\boldsymbol{x}$(and $\bar{\boldsymbol{x}}$) and $\boldsymbol{\alpha}$, we can use variation in $\bar{\boldsymbol{x}}$ (across units) to control for unobserved α_i , hence, we write the following linear (in the parameters) auxiliary relation:

$$\alpha_i = \sum_{l=1}^{L} \bar{x}^l_i \delta_l + \eta^x_i,\, \forall i \tag{14}$$

where, we also assume that $\mathbb{E}\left(v_i | \, \bar{\boldsymbol{x}}_{i'}\right) = 0,\, \forall i,\, i'$.[7] For $L = 1$, eq (14) yields the formulation proposed by Mundlak [1978].

[7] Alternatively, we might consider a non-parametric specification for the relationship in (14)

3.2 Unobserved heterogeneity and network interactions

Let us now move attention to the role unobserved heterogeneity may have in shaping the formation of links. Later, we are going to be more specific in describing how individual (unobserved) traits may affect the formation of ties among subjects; for the moment, we take for granted such a relationship. In this respect, if the unobserved α_i is going to affect the formation of each link g_{ijt} of which the network $g_{i\cdot t}$ of individual i is made of, is, we may write the following:

$$g_{ijt} = h\left(\alpha_i,\, v^g_{ijt}\right)\ ,\ \forall i, j, t \tag{15}$$

For the moment, we do not impose any limitation on $h\left(\right)$ and let us focus on the implications of eq (15): by hitting each single link formed by individual i, unobserved heterogeneity is going to shape each i's network and some of its underlying characteristics, hence the vector of unobserved terms $\boldsymbol{\alpha}$ correlates with the whole proximity matrix $\boldsymbol{G}$ we have encountered before. To the extent such correlation involves $\boldsymbol{GX}$ (and all its successives lags used in eq (5) as instruments for $\boldsymbol{Gy}$), assumption (9) we made before no longer holds. A similar argument is made in Qu and Lee [2015] and we report simulation evidence in this respect later.

Using the same logic we applied before, we would like to cope with $\mathbb{E}\left(g_{ijt}|\, \alpha_i\right) \neq 0$ using observed quantities which are time invariant. Indeed, this would actually imply to specify as many ancillary relations as the number of potential links each i has available[8], which - at least - may be a costly strategy in terms of parameters to be estimated.

In order to circumnvent this issue, we can re-target our focus on each individual i's network as a whole, as we have seen α_i is going to affect every link i has available. So, suppose there is some observed statistic μ^g_{it} describing features of each i's network and whose variation across units is (also) attributable to α_i. Then, we may write down an ancillary relation which mimics that we considered before (we omit the g suffix therein):

$$\alpha_i = \sum_{m=1}^{M} \bar{\mu}^m_i \psi_m + \eta^g_i,\ \forall i \tag{16}$$

where, again, we assume $\mathbb{E}\left(\eta^g_i|\, \bar{\mu}^m_{i'}\right) = 0,\ \forall i,\ i',\ m$. One remark is useful at this stage: first of all, in eq (16) we directly considered a time average of the network statistic μ_{it} as we did previously for x_{it} in order to avoid confusion, but most importantly, since averaging over T allows to better catch variation of the statistic induced by α. Nonetheless, we could consider a non averaged-out statistic as well, which makes such approach implementable also with cross-sectional data.

[8]In the current context we are not going to place an a priori restriction over the available sample, so that each observation has N-1 potential links available. Below, we are going to be more specific on this issue.

3.3 Estimation model controlling for unobserved heterogeneity

In the previous section we have written down each correction separately for exposition purposes. However, from a statistical standpoint, we are going to consider eqs (14) and (16) as one expression:

$$\alpha_i = \sum_{l=1}^{L} \bar{\boldsymbol{x}}_{\boldsymbol{i}}^{\boldsymbol{l}} \boldsymbol{\delta_l} + \sum_{m=1}^{M} \bar{\mu}_i^m \psi_m + \eta_i, \, \forall i$$

where we have maintained the previous notation as a matter of clarity. Stack across i and without loss of generality consider linear terms only:

$$\boldsymbol{\alpha} = \bar{\boldsymbol{X}}\boldsymbol{\delta} + \bar{\boldsymbol{\mu}}\psi + \boldsymbol{\eta} \tag{17}$$

We can then rewrite eq (4) using the ancillary expression above to obtain:

$$\boldsymbol{y} = \boldsymbol{G}\boldsymbol{y}\rho + \boldsymbol{X}\boldsymbol{\beta} + \bar{\boldsymbol{X}}\boldsymbol{\delta} + \bar{\boldsymbol{\mu}}\psi + \boldsymbol{\eta}\boldsymbol{\varepsilon} \tag{18}$$

and let us define $\boldsymbol{Z_{DC}} = \left(\boldsymbol{Gy}, \, \boldsymbol{X}, \, \bar{\boldsymbol{X}}, \, \bar{\boldsymbol{\mu}}\right)$, $\boldsymbol{\kappa_{2C}} = (\rho, \, \boldsymbol{\beta}, \, \boldsymbol{\delta}, \, \psi)$. Now, based on (17), let us reconsider the term $\boldsymbol{S_{G\alpha}}$:

$$\boldsymbol{S_{G\alpha}} = \sum_{l=1}^{p} \boldsymbol{G^l}\boldsymbol{\alpha}\rho^{l-1} = \sum_{l=1}^{p} \boldsymbol{G^l}\left(\bar{\boldsymbol{X}}\boldsymbol{\delta} + \bar{\boldsymbol{\mu}}\psi + \boldsymbol{\eta}\right)\rho^{l-1} \stackrel{def}{=} [\boldsymbol{S_{G\bar{X}}}, \, \boldsymbol{S_{G\bar{\mu}}}] \left[\boldsymbol{\pi}_{\bar{\boldsymbol{X}}}^{'}, \, \boldsymbol{\pi}_{\bar{\boldsymbol{\mu}}}^{'}\right]^{'} + \boldsymbol{S_{G\eta}}\boldsymbol{\rho} \tag{19}$$

where $\boldsymbol{S_{G\bar{X}}}$ and $\boldsymbol{S_{G\bar{\mu}}}$ are, respectively, $NT \times pk_1$ and $NT \times p$ matrices, while $\boldsymbol{\pi}_{\bar{\boldsymbol{X}}}$ and $\boldsymbol{\pi}_{\bar{\boldsymbol{\mu}}}$ are, respectively, $pk_1 \times 1$ and $p \times 1$ column vectors. Hence, we define $\boldsymbol{H_{DC}} = \left(\boldsymbol{S_{GX}}, \, \boldsymbol{S_{G\bar{X}}}, \, \boldsymbol{S_{G\bar{\mu}}}, \, \boldsymbol{X}, \, \bar{\boldsymbol{X}}, \, \bar{\boldsymbol{\mu}}\right)$ and, consequently $\boldsymbol{P_{DC}} = \boldsymbol{H_{DC}}\left(\boldsymbol{H}_{\boldsymbol{DC}}^{'}\boldsymbol{H_{DC}}\right)^{-1}\boldsymbol{H}_{\boldsymbol{DC}}^{'}$. Eventually, we obtain:

$$\hat{\boldsymbol{\kappa}}_{\boldsymbol{DC}} = \left(\boldsymbol{Z}_{\boldsymbol{DC}}^{'}\boldsymbol{P_{DC}}\boldsymbol{Z_{DC}}\right)^{-1}\left(\boldsymbol{Z}_{\boldsymbol{DC}}^{'}\boldsymbol{P_{DC}}\boldsymbol{y}\right) \tag{20}$$

As a matter of completeness, let us also formally define the estimator for the case unobserved heterogeneity is just relating to network ties. Actually, this means we do not need eq (14) and, accordingly, considering versions of eqs (17)-(19) without terms relating to $\bar{\boldsymbol{X}}$, we are left to define: $\boldsymbol{Z_{SC}} = (\boldsymbol{Gy}, \, \boldsymbol{X}, \, \bar{\boldsymbol{\mu}})$, $\boldsymbol{\kappa_{SC}} = (\rho, \, \boldsymbol{\beta}, \, \psi)$, $\boldsymbol{H_{SC}} = (\boldsymbol{S_{GX}}, \, \boldsymbol{S_{G\bar{\mu}}}, \, \boldsymbol{X}, \, \bar{\boldsymbol{\mu}})$ and $\boldsymbol{P_{SC}} = \boldsymbol{H_{SC}}\left(\boldsymbol{H}_{\boldsymbol{SC}}^{'}\boldsymbol{H_{SC}}\right)^{-1}\boldsymbol{H}_{\boldsymbol{SC}}^{'}$ in order to obtain:

$$\hat{\boldsymbol{\kappa}}_{\boldsymbol{SC}} = \left(\boldsymbol{Z}_{\boldsymbol{SC}}^{'}\boldsymbol{P_{SC}}\boldsymbol{Z_{SC}}\right)^{-1}\left(\boldsymbol{Z}_{\boldsymbol{SC}}^{'}\boldsymbol{P_{SC}}\boldsymbol{y}\right) \tag{21}$$

So far, we have voluntary neglected the structure of errors. However, the compound error term $\boldsymbol{u} = \boldsymbol{\alpha} \otimes \iota_T + \boldsymbol{\varepsilon}$ implies the covariance matrix is no longer diagonal, *i.e.*:

$$\boldsymbol{\Omega} = \mathbb{V}(\boldsymbol{u}) = \sigma_\alpha^2 \left(\boldsymbol{J_t} \otimes \boldsymbol{I_N}\right) + \sigma_\varepsilon^2 \boldsymbol{I_{NT}}$$

Adapting formulations used in Baltagi et al [2014], we can rearrange terms to recast terms to represent $\boldsymbol{\Omega}$ in terms of its spectral decomposition:

$$\boldsymbol{\Omega} = \lambda_1 \boldsymbol{Q_1} + \lambda_2 \boldsymbol{Q_2} \tag{22}$$

where $\lambda_1 = \sigma_\varepsilon^2$, $\lambda_2 = T\sigma_\alpha^2 + \sigma_\varepsilon^2$ and $\boldsymbol{Q}_1 = \boldsymbol{I_{NT}} - \left(\boldsymbol{I_N} \otimes \bar{\boldsymbol{J}}_{\boldsymbol{T}}\right)$, $\boldsymbol{Q}_2 = \boldsymbol{I_N} \otimes \bar{\boldsymbol{J}}_{\boldsymbol{T}}$. [9] From eq (22) we can then obtain[10]:

$$\boldsymbol{\Omega^{-1}} = \lambda_1^{-1} \boldsymbol{Q_1} + \lambda_2^{-1} \boldsymbol{Q_2} \tag{23}$$

Premultiplication of both sides of (18) by $\boldsymbol{\Omega^{-1/2}}$ yields a GLS-SC2SLS estimator:

$$\hat{\boldsymbol{\kappa}}\boldsymbol{*_{DC}} = \left(\boldsymbol{Z'_{DC} P^*_{DC} Z_{DC}}\right)^{-1} \left(\boldsymbol{Z'_{DC} P^*_{DC} y}\right) \tag{24}$$

or, in the case we were only correcting for the endogenoues ajdacency matrix:

$$\hat{\boldsymbol{\kappa}}\boldsymbol{*_{SC}} = \left(\boldsymbol{Z'_{SC} P^*_{SC} Z_{SC}}\right)^{-1} \left(\boldsymbol{Z'_{SC} P^*_{SC} y}\right) \tag{25}$$

with $\boldsymbol{P^*_{SC}} = \boldsymbol{H}_{(\cdot)\boldsymbol{C}} \left(\boldsymbol{H}'_{(\cdot)\boldsymbol{C}} \boldsymbol{\Omega} \boldsymbol{H}_{(\cdot)\boldsymbol{C}}\right)^{\boldsymbol{-1}} \boldsymbol{H}'_{(\cdot)\boldsymbol{C}}$, $(\cdot) = D, S$. With unknown parameters to be estimated enter the expression for $\boldsymbol{\Omega}$, the feasible-GLS estimation of equations (24) and (25) would yield a panel version of the Keleijan and Prucha [1998] estimator.

4 Network formation with unobserved heterogeneity

In this section we provide a simple example in which observation-level unobserved heterogeneity affects the network formation process. We wish to use a link generating function displaying a flexible shape as concerning its argument, where changes in its shape would be the product of individual heterongeity.

Let us define as sender that node i whom network (and underlying links) we are focusing on and let us define as receiver every node j the sender can choose as component of his network. The weight that each link between any sender i and each receiver j has on i's network is given by the following:

$$g_{ijt} = \frac{exp\left(\gamma_0 + \gamma_1 \alpha_i z_{jt} + \gamma_2 c_{ij} + v_{ijt}\right)}{\sum_j exp\left(\gamma_0 + \gamma_1 \alpha_i z_{jt} + \gamma_2 c_{ij} + v_{ijt}\right)}, j \neq i \tag{26}$$

First of all, we set $g_{ijt} = 0$, $\forall i = j$, t - that is we rule out reflexive ties - and note that weights formed according to eq (26) are not necessarily symmetric or, put another way, we are focusing on a directed network.

[9] $\bar{\boldsymbol{J}}_{\boldsymbol{N}} = \frac{\boldsymbol{J_N}}{N}$ and the two matrices at the numerator are matrices of ones.

[10] See also Wansbeek Kapteyn [1982].

Links formed according to eq (26) obey a normalized exponential function - a.k.a. softmax function - the range of this being bounded within the $[0,1]$ interval and such that $\sum_j g_{ijt} = 1$. For low values of the numerator relative to the denominator $g_{ijt} \to 0$, which is a convenient approximation to a setup in which units do not establish ties with the whole set of potential nodes.[11] Network weights formed this way are consistent with a setup in which "choosing" units derive (some notion of) utility from every alternative j given by the argument entering the exponential term in the aformentioned equation, Eventually, the resulting g_{ijt} may invariantly be seen as the probability individual j has to be chosen by individual i at time t or the weight the former is having on i's network. However, the latter statement implies that link-formation and weight attribution occur simultaneously.

The tie formation mechanism reported above relates to a wide amount of empirical research in industrial organization[12] and to studies accounting for spillovers effects, like in Bayer and Timmins [2007].[13]

We attribute an a priori meaning to each term entering the argument of the logistic function. For instance, $\alpha_i \in (0,1)$ and reflects observational heterogeneity in evaluating the true z_j, irrispectively of the time period. For instance, such heterogeneity could represent individual propensities to evaluate potential peers or just reflect heterogenous preferences. The J dimensional vector $\boldsymbol{z}$ can be thought as some kind of "selection" vector potentially available to any observational unit when deciding about establishing a tie. It is worth stressing that z_j must exhibit some degree of variability across observations as well; intuitively, if candidates peers would not be heterogenous among themselves, it had been useless for each i to have any such "selection skill" , hence variation in α_i would turn uneffective in determining weights.

The next term, c_{ij}^{rs} account for distance between any two individuals i and j belonging to any two groups r and s (where $c_{ij} = 0$ means observations do belong to the same group). We point out that, in the present context, such groups based variable is assumed to be exogenously determined and it does not enter the outcome equation (1). The presence of costs due to distance across individuals (where distance may currently be catching different concepts of distance) reduces the incentives for an individual to establish links outside its own group, other things being equal. Focusing only on α and c scalars entering the link geneating equation, the last point point holds no matter α_i, as $g\left(c_{ij}|\alpha_i\right) < g\left(c_{ij}^{'}|\alpha_i\right), \, c_{ij} < c_{ij}^{'}$. On the other hand, it does imply that conditionally on c_{ij}: $g\left(\alpha_i|c_{ij}\right) > g\left(\alpha_{i'}|c_{ij}\right), \, \alpha_i > \alpha_{i'}$.

The parameters γ_1 and γ_2 account for the size the respective terms have in determining weights, while γ_0 may be interpreted as an overall propensity to establish (or not) ties. Finally, v_{ijt} is comprehensive of any other receiver-sender-time factor impacting on the corresponding link.

[11] As a robustness check, we have conducted simulation exercises in which all those $g_{ijt} < g_{ijt}^{*} = .005$ are set to zero. As expected, results do not change significantly as compared to those in which we do not apply such threshold.

[12] In particular, we refer to the so-called "characteristics space" approach to demand estimation - see, for instance, Berry [1994] and Berry, Levinsohn and Pakes [1995].

[13] More generally, Blume [1993], Brock [1993], Durlauf (1993), Brock and Durlauf [1993] were among the most important contributions in introducing binary choice models in the context of social interactions. See also Blume et al. [2011].

The next three figures highlight some features of indvidual networks resulting from eq (26) and the implied network structure. Figure 1 plots indivual network weights as a function of z conditional on individual type α for three different values of the parameter γ_1 and set of potential verteces $J = 100$. In order to single out the role played by unobserved heterogeneity, we do not consider the role of the remainder term v_{ijt} and set t=1.

The four quadrants of the plot highlight how an individual network is shaped by α. In the first quadrant $g_{ij} \to \bar{g} = \frac{1}{J}$, which implies i's outdegree approaches the maximal outdegree (which also implies full connectivity); furthermore, peers' selection variable z plays no role in i's arcs formation. Moving across the other quadrants, selectivity in the formation of ties progressively arises and, in this respect, we can also appreciate the different magnitude associated with different parameters values.

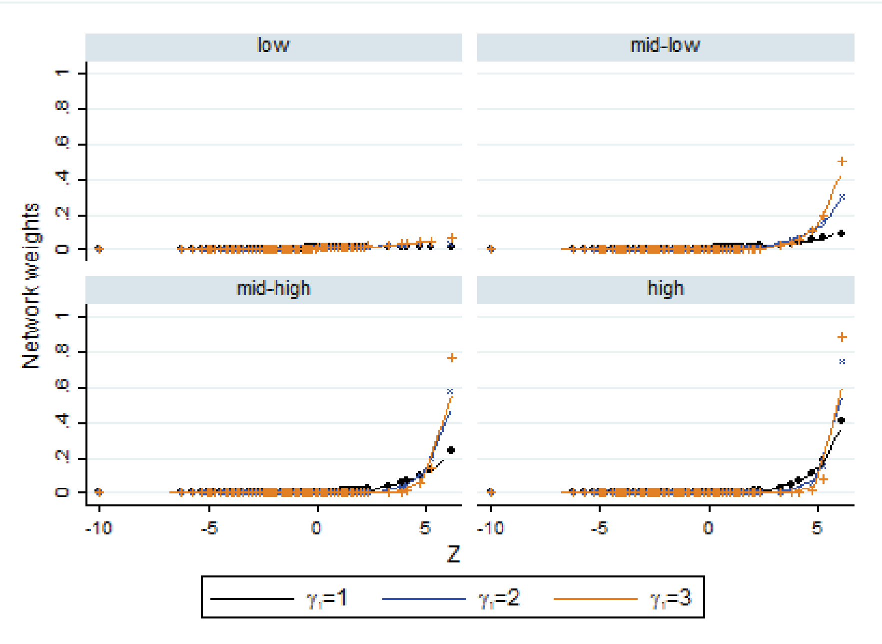

Figure 1: Network weights and Z by α type

Table 1 reports summary statistics on network characteristics for the network of senders. First, it turns out no unit is isolated, while $d_O^{max} = \frac{1}{2}J$. Also, the maximal weigth attributed to a tie is close to .9, which means such unit has formed an almost exclusive tie.

Table 1: Summary statistics for individual networks

Variable	Obs	Mean	Std. Dev.	Min	Max
Outdegree	50	17.58	12.99	4	57
Range	50	0.44	0.21	0.07	0.84
Std. Dev.	50	0.05	0.02	0.01	0.08

Based on I=50 senders and J=100 receivers ($\gamma_1 = 2, \gamma_2 = -2$)

Based of the same setting of previous graphs, Figure 2 plots weights frequencies conditionally on α. The mass of frequencies gravitate around $\bar{g} = .01$, though the vertical bar slims down for highly α endowed individuals - meaning that a greater number of arcs are indeed zero.[14]

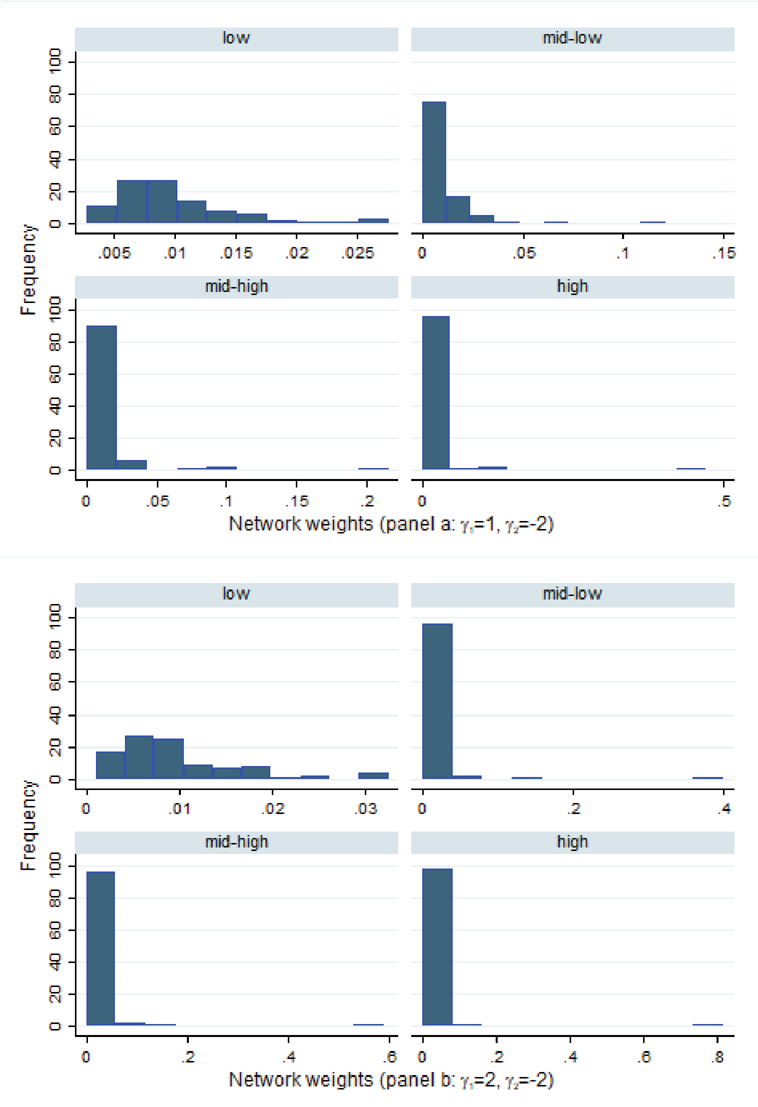

Figure 2: Network weights densities by α type

Pronounced selectivity emerges across panels for "high" αs, though maximal arcs size gets bigger for higher values of the parameter γ_1 (see panel c below).

[14]Recall that we conventionally set to zero weights $g_{ij} < 0.005$. The number of zero weights for "low" α is 12, 25 and 38 for $\gamma_1 = 1, 2, 3$ respectively.

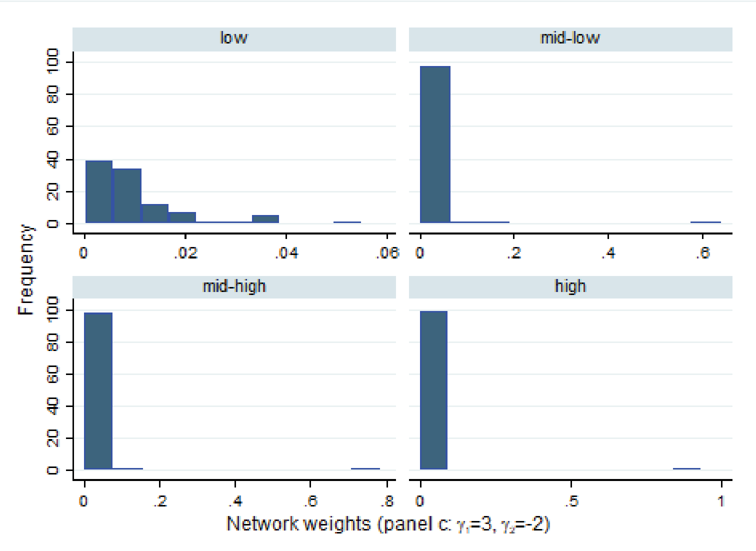

Figure 2': Network densities by α type (cont'd)

Finally, in the three panels of Figure 3 we plot the outdegree frequency distributions associated with the already mentioned three values of γ_1 . The (out)degree distribution is a crucial statistic for describing the structure of a network. For instance, it does embed information concerning connectedness patterns of a network as well as the shape of heterogeneity within a network; further it may result a sufficient statistics for the determinants of network formation, like in Graham [2017].[15]

We actually consider J=50 realizations and include the remainder term v_{ijt}in determining weights allowing for time variation. The most important message provided by the three panels below concerns the rise of network level concentration for outdegree cardinality as concerning higher values of γ_1. That is, the whole network displays substantial stability over the support, though increasing γ_1 reduces the "counfounding" role played by v_{ijt}. Interestingly, moving across panels, the outdegree distribution seems to be somewhat mixing between the so-called "scale-free" distribution and that of a Poisson network.[16] Moreover, highest nodes cardinalities gravitate far away from maximal outdegree as well as non-reiceving (i.e. isolated) nodes are avoided. In the context of dichotomous and/or symmetric networks, it has been shown that completeness (i.e. extreme connectedness) may harm identification of social effects.[17] Although, this is less a concern in the current directed and row-normalized network, completeness may still cause some unpleasant consequences for identification, as well as not being a prominent feature of many real world networks.

[15]See also Wasserman and Faust [1994] and Graham [2015].
[16]see Jackson [2008].
[17]See, for instance, Bramoullé et al. [2009] or Blume et al. [2011] .

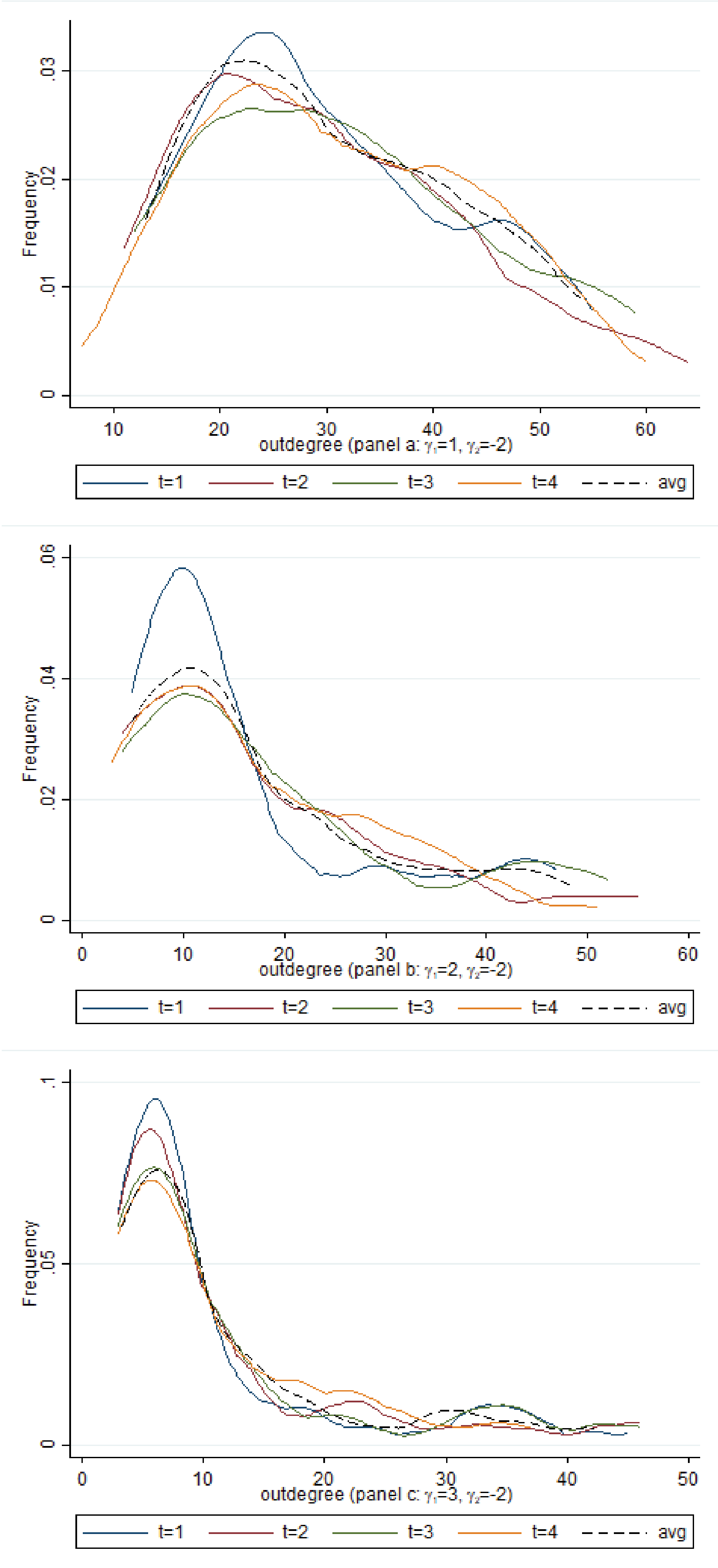

Figure 3: Outdegree densities

4.1 Network statistics and unobserved heterogeneity

Now that we have described some features of the network structure in the discussion above, we present stylized evidence regarding the link between endogenoues $\boldsymbol{\alpha}$ and various network statistics considered at the individual level. It is not immediate to apply some of the commonly used statistics describing network architecture in the current environment, for we also interested in statistiics displaying variation over units. However, one of the salient features emerging from the previous discussion is the extent to which heterogeneity affects the network structure. On this ground, we are going to focus attention onto three statistics and their relation with α: the range, the standard deviation and outdegree of individual networks, being defined, respectively, as:

$$R_{G_{i.t}} = max\left(g_{ijt}\right) - min\left(g_{ijt}\right) \approx max\left(g_{ijt}\right), \; j \neq i \tag{27}$$

$$\sigma_{G_{i.t}} = \sqrt{\frac{1}{J}\sum_{j \neq i}\left(g_{ijt} - \frac{1}{J}\right)^2} \tag{28}$$

$$d_{O,G_{i.t}} = \sum_{j=1}^{J} \mathit{1}\left(g_{ijt} > 0\right) \tag{29}$$

where, in the current setting: $J = N - 1$, and set an approximation threshold for the indicator defining the outdegree function $g^{*}_{ijt} = 0.005$.

Figg. 4-6 document the relationship between unobserved heterogeneity and the network indicators introduced above over various combinations of the parameters γ_1 and γ_2. To be precise, in the following figures we plot time-averages of those statistics, as they enter the estimator we defined before.

There are two salient stylized facts showing up: all the statistics considered are strongly associated with individual unobserved heterogeneity, and this is true irrespectively of the parameters combination used. The second aspect to consider concerns the shape of the relationship, which may change between and within statistic (especially for the outdegree measure) for different combination of parameters. The latter aspect is more evident by contrasting Figure 6 with the precedings: for both range and standard deviation of individual networks there seem to emerge a convex relationship with α. It is worth noting that such non-linearity would not being picked-up by a simple correlation measure and, most importantly, allows to underline that a non-linear functional approximation for $\mathbb{E}\left(\boldsymbol{\alpha}|\,\bar{\boldsymbol{\mu}}\right)$ should be considered when implementing the correction proposed earlier.

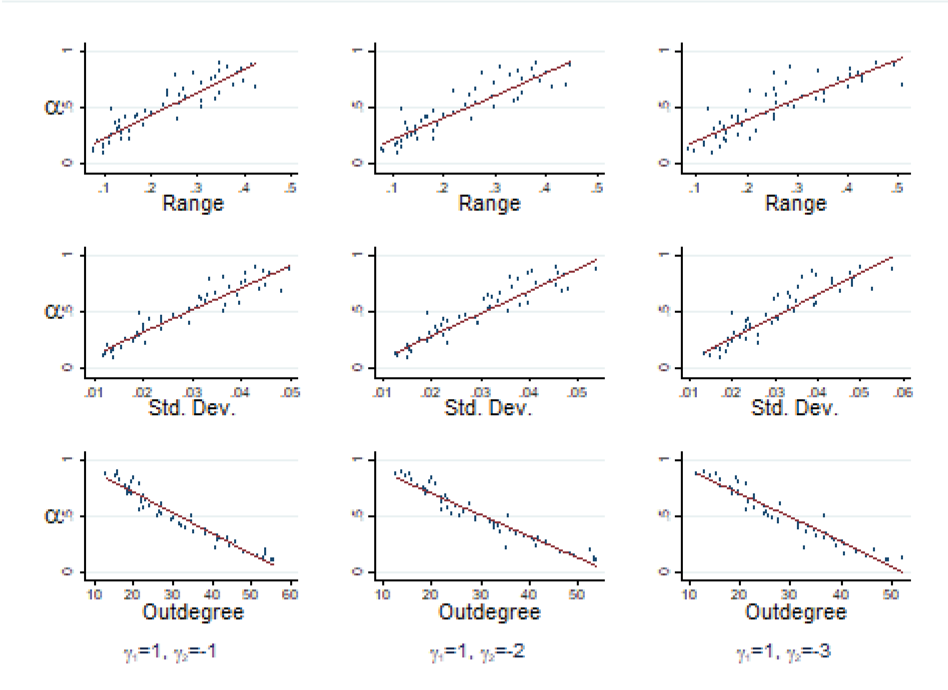

Figure 4: Unobserved heterogeneity and network stats (1)

In order to strengthen the argument we are making, let us consider the partial derivative of $R_{Gi\cdot t}$ with respect to α. To simplify things, let us restrict attention to the term entailing α in the numerator of eq (26) and focus on $max\left(g_{ijt}\right)$ in eq (27), leaving aside the time index, set $\gamma_1 = 1$ and $z^* : \; g_{ijt}\left(z^*\right) \geq g_{ijt}\left(z_j\right), \forall j$, so that :

$$\frac{\partial max\left(g_{ijt}\right)}{\partial \alpha_i} = \frac{exp\left(\gamma_1 \alpha_i z^*\right)}{\sum\limits_{j\neq i} exp\left(\gamma_1 \alpha_i z_j\right)} =$$

$$= \frac{z^* exp\left(\alpha_i z^*\right)\sum_{j\neq i} exp\left(\alpha_i z_j\right) - exp\left(\alpha_i z^*\right)\sum_{j\neq i} z_j exp\left(\alpha_i z_j\right)}{\left[\sum\limits_{j\neq i} exp\left(\gamma_1 \alpha_i z_j\right)\right]^2} =$$

$$= \frac{exp\left(\alpha_i z^*\right)}{\left[\sum\limits_{j\neq i} exp\left(\gamma_1 \alpha_i z_j\right)\right]^2}\left(z^* \sum_{j\neq i} exp\left(\alpha_i z_j\right) - \sum_{j\neq i} z_j exp\left(\alpha_i z_j\right)\right) \propto$$

$$\propto z^* \sum_{j\neq i} exp\left(\alpha_i z_j\right) - \sum_{j\neq i} z_j exp\left(\alpha_i z_j\right) =$$

$$= \sum_{j\neq i}\left(z^* - z_j\right) exp\left(\alpha_i z_j\right) > 0\,, \;\; \forall z_j < z^*$$

Since, $z^* \geq z_j, \forall j$, the result holds with strict inequality unless $z^* = z_j$. Using the same argument - reversed - it turns out that $\frac{\partial min(g_{ijt})}{\partial \alpha_i} > 0$, which implies that $\frac{\partial R_{Gi\cdot t}}{\partial \alpha_i} > 0$ unless $max\left(g_{ijt}\right) = min\left(g_{ijt}\right)$.[18]

[18]Indeed, the way links are actually formed generally produces zero weights, so that approximating $R_{Gi\cdot t} = max\left(g_{ijt}\right)$ is perfectly fine.

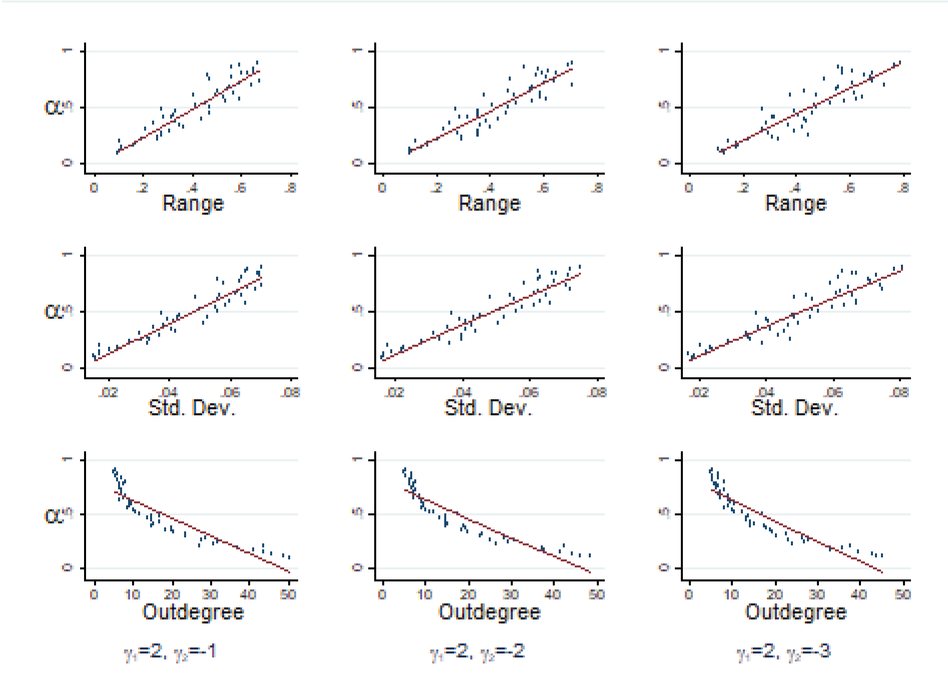

Figure 5: Unobserved heterogeneity and network stats (2)

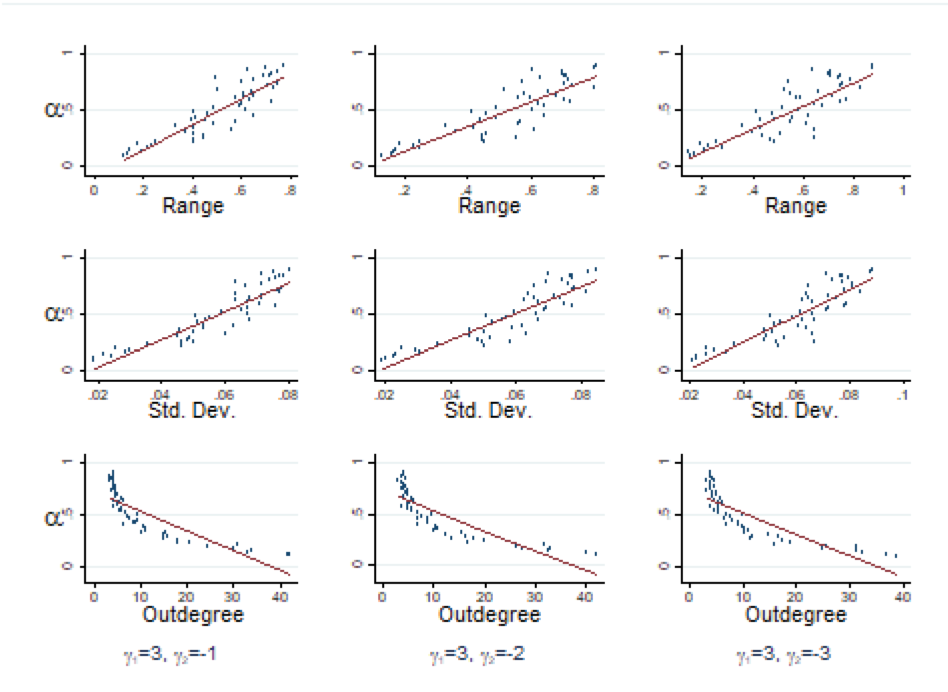

Figure 6: Unobserved heterogeneity and network stats (3)

We anticipate that in our simulation exercises we are going to rely on the time averages of the range and standard deviation to proxy unobserved heterogeneity. Despite a seemingly tight link with α, we leave aside outdegree measures, mainly because such measure may be mis-interpreted with respect to the underlying mechanism associated with α. That is, one may come to the conclusion that heterogeneity simply impacts on the cardinality of each i's network; though the connectedness of an individual's network is surely affected by higher or lower realizations of α, the latter is going to affect the strength of ties, which is given by the size of weights. Nonetheless, based on the heuristical evidence reported above, (out)degree is a valid candidate proxy as well.

5 Simulation experiment

5.1 Data generating process and setup description

In what follows we report simulation evidence for the corrections proposed before. We break down results for the case in which unobserved heterogeneity only affects network formation and those in which also exogenous effects correlate with unobservables. For both set of results we consider different combinations of the network parameters, namely: $(\gamma_1, \gamma_2) = \{(1,\, 2,\, 3) \times (-1,\, -2,\, -3)\}$. Moreover, we perform various simulation exercises under the scenario in which only network interactions are affected by unobservables, that is, we consider i) results using non-linear (second-order polynomial) corrections ; ii) we investigate the behaviour of the proposed correction over two different sizes for the groups defining the distance measure ,c_{ij}, $N_C = (10,\, 20)$; iii) we do compare the models with both linear and non-linear correction with fixed-effects and first differencing estimators; lastly, we repeat our simulation exercise for two different values of the parameter defining endogenous effects, $\rho = (.3,\, .7)$[19].

The data generating process is $\boldsymbol{Y} = [\boldsymbol{I} - \rho\boldsymbol{G}]^{-1}(\boldsymbol{X}\beta_1 + \boldsymbol{u})$, where the matrix of exogenous effects is actually restricted to be the $NT \times 1$ vector $\boldsymbol{x}_1$ following an $AR(1)$ process: $x_{1it} = \phi_x x_{1it-1} + \varepsilon^x_{it}$, $\forall i, t$, with $\phi_x = 0.8$. We consider x_{1i0} and ε^x_{it} as *i.i.d* draws from a standard normal distribution. The regression parameter associated with exogenous effects $\beta_1 = .4$. In the unobserved part of the model,$\boldsymbol{u} = \boldsymbol{\alpha} \otimes \boldsymbol{i}_T + \boldsymbol{\varepsilon}$, $\boldsymbol{\alpha}$ is a $N \times 1$ vector of i.i.d draws from a standard uniform distribution with support $(0, 1)$, while the $NT \times 1$ vector $\boldsymbol{\varepsilon} \sim N(0,\, \boldsymbol{I_{NT}})$; we rule out correlation between unobservables, hence $\mathbb{E}(\alpha_i,\, \varepsilon_{it}) = 0$, $\forall i, t$.

In eq (26) we set $\gamma_0 = 0$ and consider the triplets of values described above for $\gamma_1$and γ_2. We have already defined $\boldsymbol{\alpha}$, while we actually consider $z_{jt} = z_j \overset{i.i.d.}{\sim} N(1,\, 3)$, while v_{ijt} are draws from a standard normal distribution.[20]

Finally, the variable accounting for extra-group distance is defined as: $c^{rs}_{ij} = |\theta^r_i - \theta^s_j|$, and each $\theta^{(\cdot)}_i$ derives from randomly assigned group membership, based on draws from a uniform distribution with support $\in (0, 1)$.

All the results reported are based on 1000 simulations and all the tables restrict to statistics concerning ρ and β_1. Nominal coverage is set to .95.

[19]We comment in the main text results based on $N_C = 10$ only. In Appendix 2, we report results based on setting the expected group numerosity to 20; we do so by i) keeping the number of groups fixed to 10, while increasing cross-sectional units to N=200 and ii) by keeping N=100 and reducing the number of groups to G=5.

[20]We have also conducted some preliminary sensitivity analysis over $\boldsymbol{z}$, considering various combinations of parameters for the mean and standard deviation and results appear to be stable across those combinations. The same seems to be true when considering an $AR(1)$ process for z_{jt}.

5.2 Simulation results

We present a preview of results in the two panels of Figure 7. We consider the empirical densities of the estimators involving the use of network statistics as proxy for unobserved heterogeneity, namely, network range and network standard deviation. In both panels network parameters are the following: $\gamma_1 = 2$ and $\gamma_2 = -2$. The two panels differ as concerning the nominal value of ρ (the dashed vertical line in figures).

The density of the estimator based on network range is relatively more centered around the true vale and also displays lower probabilities in the tails and this is the case in both panels. Indeed, if we compare densities across panels, we also note that for higher ρ the empirical distributions of both estimators are even more precise and centerd around the true value.

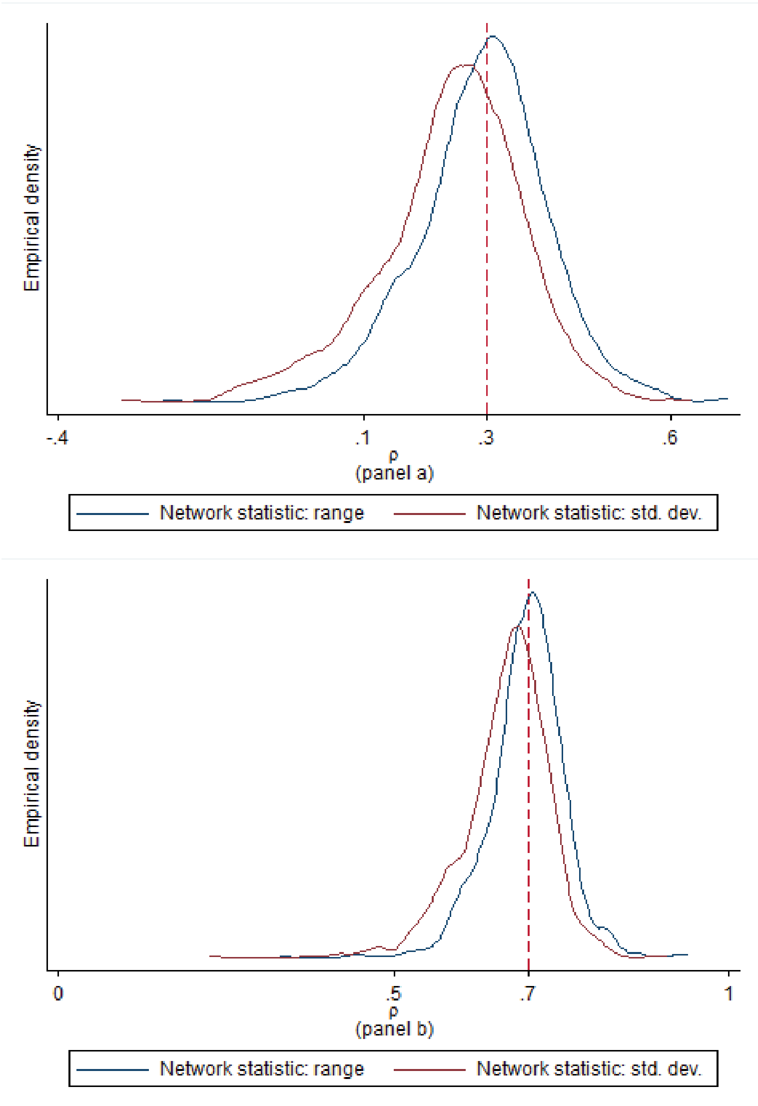

Figure 7: Empirical densities of estimators

Tables (2)-(10) report simulation results for models estimated using linear versions of the averages of network range and standard deviation. For each combination of parameters and correction we estimate four models, according to the lags used. Results concern the empirical mean, standard deviation and coverage of each estimator over the simulation sample. Results can be summarized as follows:

1. Statistics for β_1 display extreme stability across all estimators and models, with $\hat{\beta}_1 = .4$, $\hat{s.e.}_{\hat{\beta}_1} \approx .04$ and empirical coverage approaching the nominal one. The only exception is given by the models using just one lag in the instrument matrix for the estimator using no correction for unobserved heterogeneity.

2. The estimator for ρ which uses no correction for unobserved heterogeneity displays substantial bias, while estimated standard errors are generally higher as compared with the estimators employing some correction and empirical coverages are really low. There does not seem to emerge any pattern related to the combination of network parameters as concerning estimated standard errors, though the bias is higher for $\gamma_1 = 1$. Again, models using just one lag in the instrument exhibit substantial instability.

3. Comparing statistics for the two correction proposed, the estimator using the network range prevails on that using network standard deviation over all statistics considered. In particular, estimated standard errors and empirical coverage are never, respecevitly, higher and lower for the former and, as concerning coverage, the estimator considering the range lies around .9 only for $\gamma_1 = 1$, then approaching nominal coverage. The estimator using network standard deviation (henceforth n.s.d.) is relatively more biased than that using the range, though represents a sensible improvement with respect to the estimator considering no correction, the bias lying around (.09, ..06). Indeed, the bias of the range estimator is really small for those parameters combinations in which $\gamma_1 = 2$ and lies around $+/-.03$ for, respectively, $\gamma_1 = 1$ and $\gamma_1 = 3$. Further, results are substantially stables across models using different lags and, somewhat surprisingly, also models using just one lag perform well over the statistics considered.

Motivated by the stylized evidence of the previous section, in Tables (11)-(13) we report results from using 2nd order polynomial corrections for both range and n.s.d. for combination of parameters involving $\gamma_1 = 3$. The performances of the estimators based on both corrections sensibly improve (the improvement is especially notable for the n.s.d. estimator), even though standard errors are sligthly higher.

We now make an attempt to compare both linear and polyonomail range and n.s.d. estimators with fixed-effects (wìthin) and first-differencing estimators which, indeed, are generally used to cope with correlated individual and time-invariant effects[21]. We use the same settings as of the previous results and focus models using two and four lags in the instrument matrix, only to avoid proliferation of results. Tables (14)-(22) report the results, where, the first four columns report mean squared errors (MSE), while the fifth to the eighth column report empirical mean as reference. Rows refer to estimators. We try to summarize results as follows:

[21] See, for instance, Wooldridge [2010]

1. MSE for β_1 is always close or equal to zero, no matter the estimator, the model of the combinations of network paramaters.

2. MSE of both linear and non-linear correction estimators are almost alway close to zero for $\gamma_1 = 2, 3$ irrespectively of the number of lags used. For $\gamma_1 = 1$, linear corrections estimators perform better in a MSE sense than the non-linear counterparts[22], while the latters have better performances for models using four lags. Nonetheless, even in those cases, they always outperform in a MSE sense either fe and fd estimators, sometimes in a non-negligible manner.

3. Fe and fd estimators generally display lower MSE in models using four lags and it is also generally true that fe estimator performs better than fd estimator. For models using two lags, both estimators lie around .1 in the best case scenario, which is for combinations of network parameters fixing $\gamma_1 = 3$. Analogously, as concerning four lags models, fe and fd estimators get really close to the set of network corrections estimators only for $\gamma_1 = 3$.

In Tables 23-31 we repeat the same comparison exercise for $\rho = .7$. So, the purpose of those results is twofold: on one side we assess the performance of the network corrections estimators for stronger effects associated to the network multiplier; at the same time we evaluate those estimator against fe/fd counterparts once more. The previous results are confirmed and reinforced. In particular: network corrections estimators are everywhere equal or really close to zero, while fe/fd start getting close to zero when the model is estimated using four lags and $\gamma_1 = 2, 3$, while they almost never get close to zero when we use a two lags model.

Finally, Tables 32-40 report MSE and empirical mean for the parameters of interest under the scenario in which unobserved heterogenous components also correlate with $\boldsymbol{x_1}$. We set again $\rho = .3$. Rows 1 and 2 refer to estimator which, respectively, do not correct neither source of endogeneity and correct for endogenoues $\boldsymbol{x_1}$ only. Our main concern is actually on β_1; in this respect the results are satisfactory in that the correction is effective, no matter the model or the combination of network parameters. Moving attention to the parameters, we can appreciate that with no correction, β_1 is upward biased. Lastly, we note that this further correction also benefits ρ in a MSE sense, as we can appreciate by looking the rows considering network corrections (this time paired with the correction using time average of $\boldsymbol{x_1}$).

6 Concluding remarks

We propose an approach to estimate models with network interactions in the presence of individual unobserved heterogeneity. The latter may impact the formation of ties and/or exogenous effects, thereby undermining identification of the associated parameters. In a panel setting, we

[22]Again, this is in line with the anecdotal evidence drawn before.

device a way to cope with these sources of endogeneity by relying on observable variations. When exogenous effects are concerned, one can control for unobserved heterogeneity by including time-averages of the endogenous variables in the systematic part of the model. When unobserved individual traits affect the process of network formation, it is possible to explore the role of network statistics affected by those unobserved traits. This is desirable and probably unavoidable, at least from a dimensionality standpoint. As we have seen, if one would like to manage each link as standalone, this implied to device as many corrections - or, any other remedy - as the number of potential links. This is pointless, as long as one has an easy and effective alternative at hand.

Such possibility seems to be not fully explored yet. Still, the network literature has emphasized the importance of network statistics, though it is actually crucial to focus on statistics displaying variation over individuals. We derive a 2SLS estimator where where we address simultaneity bias using instruments provided by the product between successive powers of the network matrix and the matrix of exogenous covariates and then complement the identification strategy with the control approach highlighted above in order to cope with unobserved heterogeneity. Identification issues are always a concern when using instruments based on the network matrix, although in the current context it is less so, because of some intrinsic properties of the network structure. In this respect, a full derivation of a model of network formation is beyond the scope of this work; indeed, we have reported some descriptive analysis documenting insights about the "behavioural" aspects of the implied network structure and the effective role of network statistics in explaining unobserved variation. This approach can be easily extended to accomodate more general error structures - including, for instance, group unobservables.

The method we propose is also easy to implement as long as one has information on the structure of the relevant network. In this respect, it is worth stressing that, in principle, one does not need to have full access to entire structure of links, as availability of the aforementioned statistics would suffice as long as one has some (consistent) prior on the relevant network.

In a panel setting, we assess the performances of our method via a Monte Carlo exercise, in which we consider various combination of models and intensities of network interactions. We further account for different intensities of the social multiplier and separately assess the cases in which unobserved sources hit the network structure only or act on exogenous effects as well. Focusing on the former case, we underline that our approach may be also applied when a simple cross-section is available, which makes it even more valuable under data constraints.

References

Anselin, L. *"Spatial Econometrics: Methods and Models"*. Kluwer Academic Publishers, Boston, 1988.

Baltagi, B. and Fingleton, B. and Pirotte, A. "Spatial lag models with nested random effects: An instrumental variable procedure with an application to English house prices". *Journal of Urban Economics*, 80C:76–86, 2014.

Bayer, P. and Timmins, C. "Estimating Equilibrium Models of Sorting Across Locations". *Economic Journal*, 117(518):353–374, 2007.

Berry, S.T. "Estimating Discrete-Choice Models of Product Differentiation". *RAND Journal of Economics*, 25(2):242–262, 1994.

Berry, S.T. and Levinsohn, J. and Pakes, A. "Automobile Prices in Market Equilibrium". *Econometrica*, 63(4):841–890, 1995.

Blume, L. "The Statistical Mechanics of Strategic Interaction". *Games and Economic Behavior*, 5(3):387–424, 1993.

Blume, L. and Brock, W. and Durlauf, S. and Ioannides, Y. "Identification of Social Interactions". *Handbook of Social Economics*, 1B:853–964, 2011.

Bramoulle, Y. and Djebbari, H. and Fortin, B. "Identification of peer effects through social networks". *Journal of Econometrics*, 150(1):41 – 55, 2009.

Brock, W. "Pathways to Randomness in the Economy: Emergent Nonlinearity and Chaos in Economics and Finance". 1993. URL `http://www.santafe.edu/sfi/publications/working-papers.html`.

Cliff, A., Ord, J. *"Spatial Autocorrelation"*. Pion, London, 1973.

Corrado, L. and Fingleton, B. "Where is the Economics in Spatial Econometrics". *Journal of Regional Science*, 52(2):210–239, 2012.

Durlauf, S. "Nonergodic Economic Growth". *Review of Economic Studies*, 60:349–366, 1993.

Gibbons, S. and Overman, H.G and Patacchini, E. "Spatial Methods". *Handbook of Urban and Regional Economics*, 5:115–168, 2015.

Goldsmith-Pinkham, P. and Imbens, G. "Social Networks and the Identification of Peer Effects". *Journal of Business and Economic Statistics*, 31(3):253–264, 2013.

Graham, B.S. "Methods of Identification in Social Networks" . *Annual Review of Economics*, 7(1):465–485, 2015.

Graham, B.S. ”An econometric model of network formation with degree heterogeneity”. *Econometrica*, 85 (4):1033 – 1063, 2017.

Hausman, J. and Leonard, G. and Zona, J.G. ”Competitive Analysis with Differentiated Products”. *Annals of Economics and Statistics*, 34:143–157, 1994.

Horrace, W. and Liu, X. and Patacchini, E. ”Endogenous Network Production Function with Selectivity” . *Journal of Econometrics*, 190:222–232, 2016.

Jackson, M. O. *”Social and Economic Networks”*. Princeton, NJ: Princeton University Press, 2008.

Kelejian, H.H. and Prucha, I.R. ”A Generalized Spatial Two-Stage Least Squares Procedure for Estimating a Spatial Autoregressive Model with Autoregressive Disturbances”. *The Journal of Real Estate Finance and Economics*, 17(99), 1998.

Lee, L. ”Best Spatial Two-Stage Least Squares Estimators for a Spatial Autoregressive Model with Autoregressive Disturbances”. *Econometric Reviews*, 22 (4):307–335, 2003.

Lee, L. and Liu, X. and Patacchini, E. and Zenou, Y. ”Criminal Networks: Who is the Key Player?”. *CEPR Discussion Papers*, 8772, 2012.

Liu, X. ”Identification and Efficient Estimation of Simultaneous Equations Network Models”. *Journal of Business and Economic Statistics*, 32(4):516–536, 2014.

Manski, C. *”Identification Problems in the Social Sciences”*. Cambridge, Harvard Univ. Press, 1995.

Manski, C. F. ”Identification of endogenous social effects: the reflection problem”. *The Review of Economic Studies*, 60(3):531–554, 1993.

Mundlak, Y. ”On the Pooling of Timer Series and Cross Section Data”. *Econometrica*, 46(1): 69–85, 1978.

Paula, A., Richards-Shubik, S. and Tamer, E. ”Identifying Preferences in Networks With Bounded Degree”. *Econometrica*, 86:263–288, 2018.

Qu, X. and Lee, L. ”Estimating a spatial autoregressive model with an endogenous spatial weight matrix”. *Journal of Econometrics*, 184(2):209–232, 2015.

Wansbeek, T. and Kapteyn, A. ”A Class of Decompositions of the Variance-Covariance Matrix of a Generalized Error Components Model”. *Econometrica*, 50(3):713–724, 1982.

Wasserman, S. and Faust, K. *”Social Network Analysis”*. Cambridge: Cambridge University Press, 1994.

Wooldridge, J.M. *”Econometric Analysis of Cross Section and Panel Data”*. Cambridge, Mass.: MIT Press, 2010.

Appendix 1: Tables

1. Baseline results using linear corrections for endogenous network

Table 2: Models performances using linear corrections

Estimating Equations		ρ=.3			β_1=.4		
		no	range	sd	no	range	sd
$\boldsymbol{y} = \boldsymbol{Gy}\rho + \boldsymbol{x}\beta_1 + \boldsymbol{\alpha} + \boldsymbol{\varepsilon}$	coeff.	.38	.33	.19	.42	.40	.40
$\boldsymbol{Gy} = \boldsymbol{Gx}\phi + \boldsymbol{G\alpha} + \boldsymbol{\xi}$	std. err.	152.94	.14	.15	.82	.04	.04
	cover.	.28	.91	.86	.96	.94	.94
$\boldsymbol{y} = \boldsymbol{Gy}\rho + \boldsymbol{x}\beta_1 + \boldsymbol{\alpha} + \boldsymbol{\varepsilon}$	coeff.	.86	.34	.21	.40	.40	.40
$\boldsymbol{Gy} = \sum_{l=1}^{2} \boldsymbol{G^l x}\phi^l + \sum_{l=1}^{2} \boldsymbol{G^l \alpha}\rho^l + \boldsymbol{\xi}$	std. err.	.17	.13	.14	.05	.04	.04
	cover.	.15	.89	.88	.95	.94	.94
$\boldsymbol{y} = \boldsymbol{Gy}\rho + \boldsymbol{x}\beta_1 + \boldsymbol{\alpha} + \boldsymbol{\varepsilon}$	coeff.	.84	.33	.22	.40	.40	.40
$\boldsymbol{Gy} = \sum_{l=1}^{3} \boldsymbol{G^l x}\phi^l + \sum_{l=1}^{3} \boldsymbol{G^l \alpha}\rho^l + \boldsymbol{\xi}$	std. err.	.14	.12	.14	.04	.04	.04
	cover.	.14	.89	.88	.95	.94	.94
$\boldsymbol{y} = \boldsymbol{Gy}\rho + \boldsymbol{x}\beta_1 + \boldsymbol{\alpha} + \boldsymbol{\varepsilon}$	coeff.	.84	.33	.22	.40	.40	.40
$\boldsymbol{Gy} = \sum_{l=1}^{4} \boldsymbol{G^l x}\phi^l + \sum_{l=1}^{4} \boldsymbol{G^l \alpha}\rho^l + \boldsymbol{\xi}$	std. err.	.14	.12	.13	.04	.04	.04
	cover.	.13	.89	.88	.95	.94	.94

N=100, T=4, groups=10; values of network parameters: γ_0=0, γ_1=1, γ_2=-1; Z$\sim$ N(1,3); X$\sim$ AR(1), AR par=.8; columns refer to models where, respectively, no correction is applied, range or standard deviation are used to proxy unobserved heterogeneity.

Table 3: Models performances using linear corrections (cont'd)

Estimating Equations		ρ=.3			β_1=.4		
		no	range	sd	no	range	sd
$\boldsymbol{y} = \boldsymbol{Gy}\rho + \boldsymbol{x}\beta_1 + \boldsymbol{\alpha} + \boldsymbol{\varepsilon}$	coeff.	-.45	.33	.18	.42	.40	.40
$\boldsymbol{Gy} = \boldsymbol{Gx}\phi + \boldsymbol{G\alpha} + \boldsymbol{\xi}$	std. err.	83.92	.14	.15	1.07	.04	.04
	cover.	.31	.90	.85	.96	.94	.94
$\boldsymbol{y} = \boldsymbol{Gy}\rho + \boldsymbol{x}\beta_1 + \boldsymbol{\alpha} + \boldsymbol{\varepsilon}$	coeff.	.83	.33	.21	.40	.40	.40
$\boldsymbol{Gy} = \sum_{l=1}^{2} \boldsymbol{G^l x}\phi^l + \sum_{l=1}^{2} \boldsymbol{G^l \alpha}\rho^l + \boldsymbol{\xi}$	std. err.	.21	.13	.14	.05	.04	.04
	cover.	.16	.89	.87	.95	.94	.94
$\boldsymbol{y} = \boldsymbol{Gy}\rho + \boldsymbol{x}\beta_1 + \boldsymbol{\alpha} + \boldsymbol{\varepsilon}$	coeff.	.84	.33	.22	.40	.40	.40
$\boldsymbol{Gy} = \sum_{l=1}^{3} \boldsymbol{G^l x}\phi^l + \sum_{l=1}^{3} \boldsymbol{G^l \alpha}\rho^l + \boldsymbol{\xi}$	std. err.	.15	.13	.14	.04	.04	.04
	cover.	.14	.89	.88	.94	.94	.94
$\boldsymbol{y} = \boldsymbol{Gy}\rho + \boldsymbol{x}\beta_1 + \boldsymbol{\alpha} + \boldsymbol{\varepsilon}$	coeff.	.84	.33	.22	.40	.40	.40
$\boldsymbol{Gy} = \sum_{l=1}^{4} \boldsymbol{G^l x}\phi^l + \sum_{l=1}^{4} \boldsymbol{G^l \alpha}\rho^l + \boldsymbol{\xi}$	std. err.	.14	.12	.14	.04	.04	.04
	cover.	.14	.90	.88	.94	.94	.94

N=100, T=4, groups=10; values of network parameters: γ_0=0, γ_1=1, γ_2=-2; Z$\sim$ N(1,3); X$\sim$ AR(1), AR par=.8; columns refer to models where, respectively, no correction is applied, range or standard deviation are used to proxy unobserved heterogeneity.

Table 4: Models performances using linear corrections (cont'd)

Estimating Equations		ρ=.3			β_1=.4		
		no	range	sd	no	range	sd
$\boldsymbol{y} = \boldsymbol{Gy}\rho + \boldsymbol{x}\beta_1 + \boldsymbol{\alpha} + \boldsymbol{\varepsilon}$ $\boldsymbol{Gy} = \boldsymbol{Gx}\phi + \boldsymbol{G\alpha} + \boldsymbol{\xi}$	coeff.	-4.62	.33	.18	.47	.40	.40
	std. err.	4168.24	.14	.16	26.87	.04	.04
	cover.	.32	.90	.84	.96	.94	.94
$\boldsymbol{y} = \boldsymbol{Gy}\rho + \boldsymbol{x}\beta_1 + \boldsymbol{\alpha} + \boldsymbol{\varepsilon}$ $\boldsymbol{Gy} = \sum_{l=1}^{2} \boldsymbol{G^l x}\phi^l + \sum_{l=1}^{2} \boldsymbol{G^l \alpha}\rho^l + \boldsymbol{\xi}$	coeff.	.85	.33	.20	.40	.40	.40
	std. err.	.20	.13	.15	.05	.04	.04
	cover.	.17	.89	.86	.95	.94	.94
$\boldsymbol{y} = \boldsymbol{Gy}\rho + \boldsymbol{x}\beta_1 + \boldsymbol{\alpha} + \boldsymbol{\varepsilon}$ $\boldsymbol{Gy} = \sum_{l=1}^{3} \boldsymbol{G^l x}\phi^l + \sum_{l=1}^{3} \boldsymbol{G^l \alpha}\rho^l + \boldsymbol{\xi}$	coeff.	.83	.33	.21	.40	.40	.40
	std. err.	.15	.13	.14	.04	.04	.04
	cover.	.15	.91	.87	.94	.94	.94
$\boldsymbol{y} = \boldsymbol{Gy}\rho + \boldsymbol{x}\beta_1 + \boldsymbol{\alpha} + \boldsymbol{\varepsilon}$ $\boldsymbol{Gy} = \sum_{l=1}^{4} \boldsymbol{G^l x}\phi^l + \sum_{l=1}^{4} \boldsymbol{G^l \alpha}\rho^l + \boldsymbol{\xi}$	coeff.	.83	.32	.21	.40	.40	.40
	std. err.	.14	.13	.14	.04	.04	.04
	cover.	.14	.91	.87	.94	.94	.94

N=100, T=4, groups=10; values of network parameters: γ_0=0, γ_1=1, γ_2=-3; Z$\sim$ N(1,3); X$\sim$ AR(1), AR par=.8; columns refer to models where, respectively, no correction is applied, range or standard deviation are used to proxy unobserved heterogeneity.

Table 5: Models performances using linear corrections (cont'd)

Estimating Equations		ρ=.3			β_1=.4		
		no	range	sd	no	range	sd
$\boldsymbol{y} = \boldsymbol{Gy}\rho + \boldsymbol{x}\beta_1 + \boldsymbol{\alpha} + \boldsymbol{\varepsilon}$ $\boldsymbol{Gy} = \boldsymbol{Gx}\phi + \boldsymbol{G\alpha} + \boldsymbol{\xi}$	coeff.	.43	.29	.20	.38	.40	.40
	std. err.	11.43	.12	.13	.32	.04	.04
	cover.	.21	.93	.88	.95	.95	.94
$\boldsymbol{y} = \boldsymbol{Gy}\rho + \boldsymbol{x}\beta_1 + \boldsymbol{\alpha} + \boldsymbol{\varepsilon}$ $\boldsymbol{Gy} = \sum_{l=1}^{2} \boldsymbol{G^l x}\phi^l + \sum_{l=1}^{2} \boldsymbol{G^l \alpha}\rho^l + \boldsymbol{\xi}$	coeff.	.68	.29	.23	.40	.40	.40
	std. err.	.16	.11	.12	.05	.04	.04
	cover.	.15	.94	.90	.95	.95	.94
$\boldsymbol{y} = \boldsymbol{Gy}\rho + \boldsymbol{x}\beta_1 + \boldsymbol{\alpha} + \boldsymbol{\varepsilon}$ $\boldsymbol{Gy} = \sum_{l=1}^{3} \boldsymbol{G^l x}\phi^l + \sum_{l=1}^{3} \boldsymbol{G^l \alpha}\rho^l + \boldsymbol{\xi}$	coeff.	.66	.29	.24	.40	.40	.40
	std. err.	.13	.10	.11	.05	.04	.04
	cover.	.14	.93	.90	.95	.95	.94
$\boldsymbol{y} = \boldsymbol{Gy}\rho + \boldsymbol{x}\beta_1 + \boldsymbol{\alpha} + \boldsymbol{\varepsilon}$ $\boldsymbol{Gy} = \sum_{l=1}^{4} \boldsymbol{G^l x}\phi^l + \sum_{l=1}^{4} \boldsymbol{G^l \alpha}\rho^l + \boldsymbol{\xi}$	coeff.	.66	.29	.24	.40	.40	.40
	std. err.	.12	.10	.10	.05	.04	.04
	cover.	.14	.94	.90	.95	.95	.94

N=100, T=4, groups=10; values of network parameters: γ_0=0, γ_1=2, γ_2=-1; Z$\sim$ N(1,3); X$\sim$ AR(1), AR par=.8; columns refer to models where, respectively, no correction is applied, range or standard deviation are used to proxy unobserved heterogeneity.

Table 6: Models performances using linear corrections (cont'd)

Estimating Equations		ρ=.3			β_1=.4		
		no	range	sd	no	range	sd
$\boldsymbol{y} = \boldsymbol{G}\boldsymbol{y}\rho + \boldsymbol{x}\beta_1 + \boldsymbol{\alpha} + \boldsymbol{\varepsilon}$	coeff.	-.40	.28	.20	.60	.40	.40
$\boldsymbol{G}\boldsymbol{y} = \boldsymbol{G}\boldsymbol{x}\phi + \boldsymbol{G}\boldsymbol{\alpha} + \boldsymbol{\xi}$	std. err.	752.00	.12	.13	26.86	.04	.04
	cover.	.21	.93	.87	.95	.94	.94
$\boldsymbol{y} = \boldsymbol{G}\boldsymbol{y}\rho + \boldsymbol{x}\beta_1 + \boldsymbol{\alpha} + \boldsymbol{\varepsilon}$	coeff.	.68	.29	.23	.40	.40	.40
$\boldsymbol{G}\boldsymbol{y} = \sum_{l=1}^{2} \boldsymbol{G}^l\boldsymbol{x}\phi^l + \sum_{l=1}^{2} \boldsymbol{G}^l\boldsymbol{\alpha}\rho^l + \boldsymbol{\xi}$	std. err.	.20	.11	.12	.05	.04	.04
	cover.	.15	.93	.89	.95	.94	.94
$\boldsymbol{y} = \boldsymbol{G}\boldsymbol{y}\rho + \boldsymbol{x}\beta_1 + \boldsymbol{\alpha} + \boldsymbol{\varepsilon}$	coeff.	.67	.29	.23	.40	.40	.40
$\boldsymbol{G}\boldsymbol{y} = \sum_{l=1}^{3} \boldsymbol{G}^l\boldsymbol{x}\phi^l + \sum_{l=1}^{3} \boldsymbol{G}^l\boldsymbol{\alpha}\rho^l + \boldsymbol{\xi}$	std. err.	.13	.10	.11	.05	.04	.04
	cover.	.14	.94	.89	.95	.94	.94
$\boldsymbol{y} = \boldsymbol{G}\boldsymbol{y}\rho + \boldsymbol{x}\beta_1 + \boldsymbol{\alpha} + \boldsymbol{\varepsilon}$	coeff.	.66	.29	.24	.40	.40	.40
$\boldsymbol{G}\boldsymbol{y} = \sum_{l=1}^{4} \boldsymbol{G}^l\boldsymbol{x}\phi^l + \sum_{l=1}^{4} \boldsymbol{G}^l\boldsymbol{\alpha}\rho^l + \boldsymbol{\xi}$	std. err.	.11	.10	.10	.05	.04	.04
	cover.	.14	.94	.91	.94	.95	.94

N=100, T=4, groups=10; values of network parameters: γ_0=0, γ_1=2, γ_2=-2; Z$\sim$ N(1,3); X$\sim$ AR(1), AR par=.8; columns refer to models where, respectively, no correction is applied, range or standard deviation are used to proxy unobserved heterogeneity.

Table 7: Models performances using linear corrections (cont'd)

Estimating Equations		ρ=.3			β_1=.4		
		no	range	sd	no	range	sd
$\boldsymbol{y} = \boldsymbol{G}\boldsymbol{y}\rho + \boldsymbol{x}\beta_1 + \boldsymbol{\alpha} + \boldsymbol{\varepsilon}$	coeff.	.61	.28	.20	.37	.40	.40
$\boldsymbol{G}\boldsymbol{y} = \boldsymbol{G}\boldsymbol{x}\phi + \boldsymbol{G}\boldsymbol{\alpha} + \boldsymbol{\xi}$	std. err.	43.78	.13	.14	1.55	.04	.04
	cover.	.23	.93	.86	.95	.94	.94
$\boldsymbol{y} = \boldsymbol{G}\boldsymbol{y}\rho + \boldsymbol{x}\beta_1 + \boldsymbol{\alpha} + \boldsymbol{\varepsilon}$	coeff.	.68	.29	.22	.40	.40	.40
$\boldsymbol{G}\boldsymbol{y} = \sum_{l=1}^{2} \boldsymbol{G}^l\boldsymbol{x}\phi^l + \sum_{l=1}^{2} \boldsymbol{G}^l\boldsymbol{\alpha}\rho^l + \boldsymbol{\xi}$	std. err.	.16	.11	.12	.05	.04	.04
	cover.	.16	.94	.88	.95	.94	.94
$\boldsymbol{y} = \boldsymbol{G}\boldsymbol{y}\rho + \boldsymbol{x}\beta_1 + \boldsymbol{\alpha} + \boldsymbol{\varepsilon}$	coeff.	.67	.29	.23	.40	.40	.40
$\boldsymbol{G}\boldsymbol{y} = \sum_{l=1}^{3} \boldsymbol{G}^l\boldsymbol{x}\phi^l + \sum_{l=1}^{3} \boldsymbol{G}^l\boldsymbol{\alpha}\rho^l + \boldsymbol{\xi}$	std. err.	.13	.10	.11	.05	.04	.04
	cover.	.15	.94	.88	.94	.94	.94
$\boldsymbol{y} = \boldsymbol{G}\boldsymbol{y}\rho + \boldsymbol{x}\beta_1 + \boldsymbol{\alpha} + \boldsymbol{\varepsilon}$	coeff.	.66	.29	.24	.40	.40	.40
$\boldsymbol{G}\boldsymbol{y} = \sum_{l=1}^{4} \boldsymbol{G}^l\boldsymbol{x}\phi^l + \sum_{l=1}^{4} \boldsymbol{G}^l\boldsymbol{\alpha}\rho^l + \boldsymbol{\xi}$	std. err.	.12	.10	.11	.05	.04	.04
	cover.	.15	.94	.89	.94	.94	.94

N=100, T=4, groups=10; values of network parameters: γ_0=0, γ_1=2, γ_2=-3; Z$\sim$ N(1,3); X$\sim$ AR(1), AR par=.8; columns refer to models where, respectively, no correction is applied, range or standard deviation are used to proxy unobserved heterogeneity.

Table 8: Models performances using linear corrections (cont'd)

Estimating Equations		ρ=.3			β_1=.4		
		no	range	sd	no	range	sd
$\boldsymbol{y} = \boldsymbol{G}\boldsymbol{y}\rho + \boldsymbol{x}\beta_1 + \boldsymbol{\alpha} + \boldsymbol{\varepsilon}$ $\boldsymbol{G}\boldsymbol{y} = \boldsymbol{G}\boldsymbol{x}\phi + \boldsymbol{G}\boldsymbol{\alpha} + \boldsymbol{\xi}$	coeff.	1.21	.25	.19	.37	.40	.40
	std. err.	21.09	.12	.13	2.03	.04	.04
	cover.	.20	.92	.85	.95	.94	.94
$\boldsymbol{y} = \boldsymbol{G}\boldsymbol{y}\rho + \boldsymbol{x}\beta_1 + \boldsymbol{\alpha} + \boldsymbol{\varepsilon}$ $\boldsymbol{G}\boldsymbol{y} = \sum_{l=1}^{2} \boldsymbol{G}^{\boldsymbol{l}}\boldsymbol{x}\phi^l + \sum_{l=1}^{2} \boldsymbol{G}^{\boldsymbol{l}}\boldsymbol{\alpha}\rho^l + \boldsymbol{\xi}$	coeff.	.63	.27	.22	.40	.40	.40
	std. err.	.14	.10	.11	.05	.04	.04
	cover.	.13	.93	.90	.94	.94	.94
$\boldsymbol{y} = \boldsymbol{G}\boldsymbol{y}\rho + \boldsymbol{x}\beta_1 + \boldsymbol{\alpha} + \boldsymbol{\varepsilon}$ $\boldsymbol{G}\boldsymbol{y} = \sum_{l=1}^{3} \boldsymbol{G}^{\boldsymbol{l}}\boldsymbol{x}\phi^l + \sum_{l=1}^{3} \boldsymbol{G}^{\boldsymbol{l}}\boldsymbol{\alpha}\rho^l + \boldsymbol{\xi}$	coeff.	.61	.27	.24	.40	.40	.40
	std. err.	.11	.09	.10	.05	.04	.04
	cover.	.13	.94	.90	.95	.94	.94
$\boldsymbol{y} = \boldsymbol{G}\boldsymbol{y}\rho + \boldsymbol{x}\beta_1 + \boldsymbol{\alpha} + \boldsymbol{\varepsilon}$ $\boldsymbol{G}\boldsymbol{y} = \sum_{l=1}^{4} \boldsymbol{G}^{\boldsymbol{l}}\boldsymbol{x}\phi^l + \sum_{l=1}^{4} \boldsymbol{G}^{\boldsymbol{l}}\boldsymbol{\alpha}\rho^l + \boldsymbol{\xi}$	coeff.	.60	.27	.24	.40	.40	.40
	std. err.	.10	.09	.09	.05	.04	.04
	cover.	.14	.94	.90	.94	.94	.94

N=100, T=4, groups=10; values of network parameters: γ_0=0, $gamma_1$=3, γ_2=-1; Z$\sim$ N(1,3); X$\sim$ AR(1), AR par=.8; columns refer to models where, respectively, no correction is applied, range or standard deviation are used to proxy unobserved heterogeneity.

Table 9: Models performances using linear corrections (cont'd)

Estimating Equations		ρ=.3			β_1=.4		
		no	range	sd	no	range	sd
$\boldsymbol{y} = \boldsymbol{G}\boldsymbol{y}\rho + \boldsymbol{x}\beta_1 + \boldsymbol{\alpha} + \boldsymbol{\varepsilon}$ $\boldsymbol{G}\boldsymbol{y} = \boldsymbol{G}\boldsymbol{x}\phi + \boldsymbol{G}\boldsymbol{\alpha} + \boldsymbol{\xi}$	coeff.	-.77	.25	.19	.46	.40	.40
	std. err.	86.47	.12	.13	7.96	.04	.04
	cover.	.21	.92	.85	.95	.94	.94
$\boldsymbol{y} = \boldsymbol{G}\boldsymbol{y}\rho + \boldsymbol{x}\beta_1 + \boldsymbol{\alpha} + \boldsymbol{\varepsilon}$ $\boldsymbol{G}\boldsymbol{y} = \sum_{l=1}^{2} \boldsymbol{G}^{\boldsymbol{l}}\boldsymbol{x}\phi^l + \sum_{l=1}^{2} \boldsymbol{G}^{\boldsymbol{l}}\boldsymbol{\alpha}\rho^l + \boldsymbol{\xi}$	coeff.	.63	.27	.22	.40	.40	.40
	std. err.	.14	.10	.11	.05	.04	.04
	cover.	.15	.94	.90	.95	.94	.94
$\boldsymbol{y} = \boldsymbol{G}\boldsymbol{y}\rho + \boldsymbol{x}\beta_1 + \boldsymbol{\alpha} + \boldsymbol{\varepsilon}$ $\boldsymbol{G}\boldsymbol{y} = \sum_{l=1}^{3} \boldsymbol{G}^{\boldsymbol{l}}\boldsymbol{x}\phi^l + \sum_{l=1}^{3} \boldsymbol{G}^{\boldsymbol{l}}\boldsymbol{\alpha}\rho^l + \boldsymbol{\xi}$	coeff.	.61	.27	.23	.40	.40	.40
	std. err.	.11	.10	.10	.05	.04	.04
	cover.	.14	.94	.91	.95	.94	.94
$\boldsymbol{y} = \boldsymbol{G}\boldsymbol{y}\rho + \boldsymbol{x}\beta_1 + \boldsymbol{\alpha} + \boldsymbol{\varepsilon}$ $\boldsymbol{G}\boldsymbol{y} = \sum_{l=1}^{4} \boldsymbol{G}^{\boldsymbol{l}}\boldsymbol{x}\phi^l + \sum_{l=1}^{4} \boldsymbol{G}^{\boldsymbol{l}}\boldsymbol{\alpha}\rho^l + \boldsymbol{\xi}$	coeff.	.60	.27	.24	.40	.40	.40
	std. err.	.10	.09	.09	.05	.04	.04
	cover.	.16	.94	.91	.95	.94	.94

N=100, T=4, groups=10; values of network parameters: γ_0=0, γ_1=3, γ_2=-2; Z$\sim$ N(1,3); X$\sim$ AR(1), AR par=.8; columns refer to models where, respectively, no correction is applied, range or standard deviation are used to proxy unobserved heterogeneity.

Table 10: Models performances using linear corrections (cont'd)

Estimating Equations		ρ=.3			β_1=.4		
		no	range	sd	no	range	sd
$y = Gy\rho + x\beta_1 + \alpha + \varepsilon$ $Gy = Gx\phi + G\alpha + \xi$	coeff.	-.12	.25	.18	.39	.40	.40
	std. err.	44.05	.13	.13	1.30	.04	.04
	cover.	.22	.92	.85	.96	.94	.94
$y = Gy\rho + x\beta_1 + \alpha + \varepsilon$ $Gy = \sum_{l=1}^{2} G^l x\phi^l + \sum_{l=1}^{2} G^l \alpha\rho^l + \xi$	coeff.	.62	.27	.22	.40	.40	.40
	std. err.	.15	.10	.11	.05	.04	.04
	cover.	.16	.94	.89	.95	.94	.94
$y = Gy\rho + x\beta_1 + \alpha + \varepsilon$ $Gy = \sum_{l=1}^{3} G^l x\phi^l + \sum_{l=1}^{3} G^l \alpha\rho^l + \xi$	coeff.	.60	.27	.23	.40	.40	.40
	std. err.	.11	.10	.10	.05	.04	.04
	cover.	.15	.93	.90	.95	.94	.94
$y = Gy\rho + x\beta_1 + \alpha + \varepsilon$ $Gy = \sum_{l=1}^{4} G^l x\phi^l + \sum_{l=1}^{4} G^l \alpha\rho^l + \xi$	coeff.	.59	.27	.24	.40	.40	.40
	std. err.	.10	.09	.09	.05	.04	.04
	cover.	.16	.94	.91	.95	.94	.94

N=100, T=4, groups=10; values of network parameters: γ_0=0, γ_1=3, γ_2=-3; Z$\sim N(1,3)$; X$\sim$ AR(1), AR par=.8; columns refer to models where, respectively, no correction is applied, range or standard deviation are used to proxy unobserved heterogeneity.

2. Results using polynomial corrections

Table 11: Models performances using 2nd order polynomial corrections

Estimating Equations		ρ=.3			β_1=.4		
		no	+range	sd	no	range	sd
$y = Gy\rho + x\beta_1 + \alpha + \varepsilon$ $Gy = Gx\phi + G\alpha + \xi$	coeff.	1.21	.29	.31	.37	.40	.40
	std. err.	21.09	.18	.20	2.03	.04	.04
	cover.	.20	.93	.93	.95	.94	.94
$y = Gy\rho + x\beta_1 + \alpha + \varepsilon$ $Gy = \sum_{l=1}^{2} G^l x\phi^l + \sum_{l=1}^{2} G^l \alpha\rho^l + \xi$	coeff.	.63	.30	.30	.40	.40	.40
	std. err.	.14	.13	.14	.05	.04	.04
	cover.	.13	.93	.93	.94	.94	.94
$y = Gy\rho + x\beta_1 + \alpha + \varepsilon$ $Gy = \sum_{l=1}^{3} G^l x\phi^l + \sum_{l=1}^{3} G^l \alpha\rho^l + \xi$	coeff.	.61	.30	.30	.40	.40	.40
	std. err.	.11	.12	.12	.05	.04	.04
	cover.	.13	.94	.93	.95	.94	.94
$y = Gy\rho + x\beta_1 + \alpha + \varepsilon$ $Gy = \sum_{l=1}^{4} G^l x\phi^l + \sum_{l=1}^{4} G^l \alpha\rho^l + \xi$	coeff.	.60	.29	.30	.40	.40	.40
	std. err.	.10	.11	.11	.05	.04	.04
	cover.	.14	.93	.93	.94	.94	.94

N=100, T=4, groups=10; values of network parameters: γ_0=0, γ_1=3, γ_2=-1; Z$\sim N(1,3)$; X$\sim$ AR(1), AR par=.8; columns refer to models where, respectively, no correction is applied, range or standard deviation are used to proxy unobserved heterogeneity.

Table 12: Models performances using 2nd order polynomial corrections (cont'd)

Estimating Equations		ρ=.3			β_1=.4		
		no	range	sd	no	range	sd
$\boldsymbol{y} = \boldsymbol{G}\boldsymbol{y}\rho + \boldsymbol{x}\beta_1 + \boldsymbol{\alpha} + \boldsymbol{\varepsilon}$	coeff.	-.77	.29	.30	.46	.40	.40
$\boldsymbol{G}\boldsymbol{y} = \boldsymbol{G}\boldsymbol{x}\phi + \boldsymbol{G}\boldsymbol{\alpha} + \boldsymbol{\xi}$	std. err.	86.47	.18	.21	7.96	.04	.04
	cover.	.21	.93	.92	.95	.94	.94
$\boldsymbol{y} = \boldsymbol{G}\boldsymbol{y}\rho + \boldsymbol{x}\beta_1 + \boldsymbol{\alpha} + \boldsymbol{\varepsilon}$	coeff.	.63	.29	.30	.40	.40	.40
$\boldsymbol{G}\boldsymbol{y} = \sum_{l=1}^{2} \boldsymbol{G}^l\boldsymbol{x}\phi^l + \sum_{l=1}^{2} \boldsymbol{G}^l\boldsymbol{\alpha}\rho^l + \boldsymbol{\xi}$	std. err.	.14	.13	.14	.05	.04	.04
	cover.	.15	.93	.93	.95	.94	.94
$\boldsymbol{y} = \boldsymbol{G}\boldsymbol{y}\rho + \boldsymbol{x}\beta_1 + \boldsymbol{\alpha} + \boldsymbol{\varepsilon}$	coeff.	.61	.29	.30	.40	.40	.40
$\boldsymbol{G}\boldsymbol{y} = \sum_{l=1}^{3} \boldsymbol{G}^l\boldsymbol{x}\phi^l + \sum_{l=1}^{3} \boldsymbol{G}^l\boldsymbol{\alpha}\rho^l + \boldsymbol{\xi}$	std. err.	.11	.12	.12	.05	.04	.04
	cover.	.14	.93	.93	.95	.94	.94
$\boldsymbol{y} = \boldsymbol{G}\boldsymbol{y}\rho + \boldsymbol{x}\beta_1 + \boldsymbol{\alpha} + \boldsymbol{\varepsilon}$	coeff.	.60	.29	.30	.40	.40	.40
$\boldsymbol{G}\boldsymbol{y} = \sum_{l=1}^{4} \boldsymbol{G}^l\boldsymbol{x}\phi^l + \sum_{l=1}^{4} \boldsymbol{G}^l\boldsymbol{\alpha}\rho^l + \boldsymbol{\xi}$	std. err.	.10	.11	.11	.05	.04	.04
	cover.	.16	.94	.95	.95	.94	.94

N=100, T=4, groups=10; values of network parameters: γ_0=0, γ_1=3, γ_2=-2; Z$\sim$ N(1,3); X$\sim$ AR(1), AR par=.8; columns refer to models where, respectively, no correction is applied, range or standard deviation are used to proxy unobserved heterogeneity.

Table 13: Models performances using 2nd order polynomial corrections (cont'd)

Estimating Equations		ρ=.3			β_1=.4		
		no	range	sd	no	range	sd
$\boldsymbol{y} = \boldsymbol{G}\boldsymbol{y}\rho + \boldsymbol{x}\beta_1 + \boldsymbol{\alpha} + \boldsymbol{\varepsilon}$	coeff.	-.12	.29	.29	.39	.40	.40
$\boldsymbol{G}\boldsymbol{y} = \boldsymbol{G}\boldsymbol{x}\phi + \boldsymbol{G}\boldsymbol{\alpha} + \boldsymbol{\xi}$	std. err.	44.05	.19	.22	1.30	.04	.04
	cover.	.22	.92	.93	.96	.94	.94
$\boldsymbol{y} = \boldsymbol{G}\boldsymbol{y}\rho + \boldsymbol{x}\beta_1 + \boldsymbol{\alpha} + \boldsymbol{\varepsilon}$	coeff.	.62	.29	.29	.40	.40	.40
$\boldsymbol{G}\boldsymbol{y} = \sum_{l=1}^{2} \boldsymbol{G}^l\boldsymbol{x}\phi^l + \sum_{l=1}^{2} \boldsymbol{G}^l\boldsymbol{\alpha}\rho^l + \boldsymbol{\xi}$	std. err.	.15	.13	.14	.05	.04	.04
	cover.	.16	.93	.93	.95	.94	.94
$\boldsymbol{y} = \boldsymbol{G}\boldsymbol{y}\rho + \boldsymbol{x}\beta_1 + \boldsymbol{\alpha} + \boldsymbol{\varepsilon}$	coeff.	.60	.29	.30	.40	.40	.40
$\boldsymbol{G}\boldsymbol{y} = \sum_{l=1}^{3} \boldsymbol{G}^l\boldsymbol{x}\phi^l + \sum_{l=1}^{3} \boldsymbol{G}^l\boldsymbol{\alpha}\rho^l + \boldsymbol{\xi}$	std. err.	.11	.12	.12	.05	.04	.04
	cover.	.15	.93	.93	.95	.94	.94
$\boldsymbol{y} = \boldsymbol{G}\boldsymbol{y}\rho + \boldsymbol{x}\beta_1 + \boldsymbol{\alpha} + \boldsymbol{\varepsilon}$	coeff.	.59	.29	.29	.40	.40	.40
$\boldsymbol{G}\boldsymbol{y} = \sum_{l=1}^{4} \boldsymbol{G}^l\boldsymbol{x}\phi^l + \sum_{l=1}^{4} \boldsymbol{G}^l\boldsymbol{\alpha}\rho^l + \boldsymbol{\xi}$	std. err.	.10	.11	.11	.05	.04	.04
	cover.	.16	.94	.95	.95	.94	.94

N=100, T=4, groups=10; values of network parameters: γ_0=0, γ_1=3, γ_2=-3; Z$\sim$ N(1,3); X$\sim$ AR(1), AR par=.8; columns refer to models where, respectively, no correction is applied, range or standard deviation are used to proxy unobserved heterogeneity.

3. Comparison across corrections, within (fe) and first-differencing (fd) estimators ($\rho = .3$)

Table 14: MSE and mean of parameters for models estimated using 2 and 4 lags of interaction variables

	MSE				Mean			
	ρ (2 lags)	β_1 (2 lags)	ρ (4 lags)	β_1 (4 lags)	ρ (2 lags)	β_1 (2 lags)	ρ (4 lags)	β_1 (4 lags)
no	0.57	0.00	0.42	0.00	0.86	0.40	0.84	0.40
range	0.02	0.00	0.02	0.00	0.34	0.40	0.33	0.40
sd	0.04	0.00	0.03	0.00	0.21	0.40	0.22	0.40
range (pol.)	0.11	0.00	0.07	0.00	0.22	0.40	0.24	0.40
sd (pol.)	0.09	0.00	0.06	0.00	0.26	0.40	0.27	0.40
fe	0.74	0.00	0.09	0.00	0.39	0.40	0.34	0.40
fd	0.27	0.01	0.11	0.01	0.34	0.40	0.34	0.40

N=100, T=4, groups=10; true values of structural pars: ρ=.3 β_1=.4; true values of network parameters: γ_0=0, γ_1=1, γ_2=-1; Z$\sim$ N(1,3); X$\sim$ AR(1), AR par=.8; rows stand for: (i): no correction applied , (ii) and (iv): linear and non-linear range as proxy, (iii) and (v): linear and non-linear standard deviation as proxy, (vi) and (vii): fixed effects and first-differencing estimators

Table 15: MSE and mean of parameters for models estimated using 2 and 4 lags of interaction variables (cont'd)

	MSE				Mean			
	ρ (2 lags)	β_1 (2 lags)	ρ (4 lags)	β_1 (4 lags)	ρ (2 lags)	β_1 (2 lags)	ρ (4 lags)	β_1 (4 lags)
no	0.74	0.00	0.41	0.00	0.83	0.40	0.84	0.40
range	0.02	0.00	0.02	0.00	0.33	0.40	0.33	0.40
sd	0.04	0.00	0.03	0.00	0.21	0.40	0.22	0.40
range (pol.)	0.11	0.00	0.07	0.00	0.22	0.40	0.24	0.40
sd (pol.)	0.09	0.00	0.06	0.00	0.26	0.40	0.27	0.40
fe	0.19	0.00	0.09	0.00	0.35	0.40	0.33	0.40
fd	0.27	0.01	0.11	0.01	0.34	0.40	0.34	0.40

N=100, T=4, groups=10; true values of structural pars: ρ=.3 β_1=.4; true values of network parameters: γ_0=0, γ_1=1, γ_2=-2; Z$\sim$ N(1,3); X$\sim$ AR(1), AR par=.8; rows stand for: (i): no correction applied , (ii) and (iv): linear and non-linear range as proxy, (iii) and (v): linear and non-linear standard deviation as proxy, (vi) and (vii): fixed effects and first-differencing estimators

Table 16: MSE and mean of parameters for models estimated using 2 and 4 lags of interaction variables (cont'd)

	MSE				Mean			
	ρ (2 lags)	β_1 (2 lags)	ρ (4 lags)	β_1 (4 lags)	ρ (2 lags)	β_1 (2 lags)	ρ (4 lags)	β_1 (4 lags)
no	0.72	0.00	0.40	0.00	0.85	0.40	0.83	0.40
range	0.02	0.00	0.02	0.00	0.33	0.40	0.32	0.40
sd	0.04	0.00	0.04	0.00	0.20	0.40	0.21	0.40
range (pol.)	0.11	0.00	0.07	0.00	0.23	0.40	0.25	0.40
sd (pol.)	0.10	0.00	0.06	0.00	0.26	0.40	0.27	0.40
fe	0.15	0.00	0.09	0.00	0.34	0.40	0.33	0.40
fd	0.22	0.01	0.10	0.01	0.34	0.40	0.34	0.40

N=100, T=4, groups=10; true values of structural pars: ρ=.3 β_1=.4; true values of network parameters: γ_0=0, γ_1=1, γ_2=-3; Z∼ N(1,3); X∼ AR(1), AR par=.8; rows stand for: (i): no correction applied , (ii) and (iv): linear and non-linear range as proxy, (iii) and (v): linear and non-linear standard deviation as proxy, (vi) and (vii): fixed effects and first-differencing estimators

Table 17: MSE and mean of parameters for models estimated using 2 and 4 lags of interaction variables (cont'd)

	MSE				Mean			
	ρ (2 lags)	β_1 (2 lags)	ρ (4 lags)	β_1 (4 lags)	ρ (2 lags)	β_1 (2 lags)	ρ (4 lags)	β_1 (4 lags)
no	0.66	0.00	0.37	0.00	0.68	0.40	0.66	0.40
range	0.01	0.00	0.01	0.00	0.29	0.40	0.29	0.40
sd	0.02	0.00	0.02	0.00	0.23	0.40	0.24	0.40
range (pol.)	0.01	0.00	0.01	0.00	0.29	0.40	0.29	0.40
sd (pol.)	0.02	0.00	0.02	0.00	0.23	0.40	0.24	0.40
fe	0.13	0.00	0.05	0.00	0.31	0.40	0.32	0.40
fd	0.17	0.02	0.05	0.01	0.30	0.40	0.32	0.40

N=100, T=4, groups=10; true values of structural pars: ρ=.3 β_1=.4; true values of network parameters: γ_0=0, γ_1=2, γ_2=-1; Z∼ N(1,3); X∼ AR(1), AR par=.8; rows stand for: (i): no correction applied , (ii) and (iv): linear and non-linear range as proxy, (iii) and (v): linear and non-linear standard deviation as proxy, (vi) and (vii): fixed effects and first-differencing estimators

Table 18: MSE and mean of parameters for models estimated using 2 and 4 lags of interaction variables (cont'd)

	MSE				Mean			
	ρ (2 lags)	β_1 (2 lags)	ρ (4 lags)	β_1 (4 lags)	ρ (2 lags)	β_1 (2 lags)	ρ (4 lags)	β_1 (4 lags)
no	0.79	0.00	0.36	0.00	0.68	0.40	0.66	0.40
range	0.01	0.00	0.01	0.00	0.29	0.40	0.29	0.40
sd	0.02	0.00	0.02	0.00	0.23	0.40	0.24	0.40
range (pol.)	0.01	0.00	0.01	0.00	0.29	0.40	0.29	0.40
sd (pol.)	0.02	0.00	0.02	0.00	0.23	0.40	0.24	0.40
fe	0.19	0.00	0.05	0.00	0.30	0.40	0.32	0.40
fd	0.17	0.01	0.06	0.01	0.33	0.40	0.33	0.40

N=100, T=4, groups=10; true values of structural pars: ρ=.3 β_1=.4; true values of network parameters: γ_0=0, γ_1=2, γ_2=-2; Z$\sim$ N(1,3); X$\sim$ AR(1), AR par=.8; rows stand for: (i): no correction applied , (ii) and (iv): linear and non-linear range as proxy, (iii) and (v): linear and non-linear standard deviation as proxy, (vi) and (vii): fixed effects and first-differencing estimators

Table 19: MSE and mean of parameters for models estimated using 2 and 4 lags of interaction variables (cont'd)

	MSE				Mean			
	ρ (2 lags)	β_1 (2 lags)	ρ (4 lags)	β_1 (4 lags)	ρ (2 lags)	β_1 (2 lags)	ρ (4 lags)	β_1 (4 lags)
no	0.64	0.00	0.35	0.00	0.68	0.40	0.66	0.40
range	0.02	0.00	0.01	0.00	0.29	0.40	0.29	0.40
sd	0.03	0.00	0.02	0.00	0.22	0.40	0.24	0.40
range (pol.)	0.02	0.00	0.01	0.00	0.29	0.40	0.29	0.40
sd (pol.)	0.03	0.00	0.02	0.00	0.22	0.40	0.24	0.40
fe	0.19	0.00	0.05	0.00	0.31	0.40	0.31	0.40
fd	0.13	0.01	0.05	0.01	0.33	0.40	0.32	0.40

N=100, T=4, groups=10; true values of structural pars: ρ=.3 β_1=.4; true values of network parameters: γ_0=0, γ_1=2, γ_2=-3; Z$\sim$ N(1,3); X$\sim$ AR(1), AR par=.8; rows stand for: (i): no correction applied , (ii) and (iv): linear and non-linear range as proxy, (iii) and (v): linear and non-linear standard deviation as proxy, (vi) and (vii): fixed effects and first-differencing estimators

Table 20: MSE and mean of parameters for models estimated using 2 and 4 lags of interaction variables (cont'd)

	MSE				Mean			
	ρ (2 lags)	β_1 (2 lags)	ρ (4 lags)	β_1 (4 lags)	ρ (2 lags)	β_1 (2 lags)	ρ (4 lags)	β_1 (4 lags)
no	0.65	0.00	0.32	0.00	0.63	0.40	0.60	0.40
range	0.01	0.00	0.01	0.00	0.27	0.40	0.27	0.40
sd	0.02	0.00	0.01	0.00	0.22	0.40	0.24	0.40
range (pol.)	0.02	0.00	0.01	0.00	0.30	0.40	0.29	0.40
sd (pol.)	0.03	0.00	0.02	0.00	0.30	0.40	0.30	0.40
fe	0.08	0.00	0.03	0.00	0.31	0.40	0.31	0.40
fd	0.10	0.01	0.04	0.01	0.32	0.40	0.32	0.40

N=100, T=4, groups=10; true values of structural pars: ρ=.3 β_1=.4; true values of network parameters: γ_0=0, γ_1=3, γ_2=-1; Z$\sim N(1,3)$; X$\sim$ AR(1), AR par=.8; rows stand for: (i): no correction applied , (ii) and (iv): linear and non-linear range as proxy, (iii) and (v): linear and non-linear standard deviation as proxy, (vi) and (vii): fixed effects and first-differencing estimators

Table 21: MSE and mean of parameters for models estimated using 2 and 4 lags of interaction variables (cont'd)

	MSE				Mean			
	ρ (2 lags)	β_1 (2 lags)	ρ (4 lags)	β_1 (4 lags)	ρ (2 lags)	β_1 (2 lags)	ρ (4 lags)	β_1 (4 lags)
no	0.66	0.00	0.32	0.00	0.63	0.40	0.60	0.40
range	0.01	0.00	0.01	0.00	0.27	0.40	0.27	0.40
sd	0.02	0.00	0.01	0.00	0.22	0.40	0.24	0.40
range (pol.)	0.03	0.00	0.01	0.00	0.29	0.40	0.29	0.40
sd (pol.)	0.03	0.00	0.02	0.00	0.30	0.40	0.30	0.40
fe	0.09	0.00	0.03	0.00	0.32	0.40	0.31	0.40
fd	0.10	0.01	0.04	0.01	0.32	0.40	0.32	0.40

N=100, T=4, groups=10; true values of structural pars: ρ=.3 β_1=.4; true values of network parameters: γ_0=0, γ_1=3, γ_2=-2; Z$\sim N(1,3)$; X$\sim$ AR(1), AR par=.8; rows stand for: (i): no correction applied , (ii) and (iv): linear and non-linear range as proxy, (iii) and (v): linear and non-linear standard deviation as proxy, (vi) and (vii): fixed effects and first-differencing estimators

Table 22: MSE and mean of parameters for models estimated using 2 and 4 lags of interaction variables (cont'd)

	MSE				Mean			
	ρ (2 lags)	β_1 (2 lags)	ρ (4 lags)	β_1 (4 lags)	ρ (2 lags)	β_1 (2 lags)	ρ (4 lags)	β_1 (4 lags)
no	0.71	0.00	0.32	0.00	0.62	0.40	0.59	0.40
range	0.01	0.00	0.01	0.00	0.27	0.40	0.27	0.40
sd	0.02	0.00	0.02	0.00	0.22	0.40	0.24	0.40
range (pol.)	0.03	0.00	0.02	0.00	0.29	0.40	0.29	0.40
sd (pol.)	0.03	0.00	0.02	0.00	0.29	0.40	0.29	0.40
fe	0.08	0.00	0.03	0.00	0.32	0.40	0.31	0.40
fd	0.12	0.01	0.04	0.01	0.33	0.40	0.32	0.40

N=100, T=4, groups=10; true values of structural pars: ρ=.3 β_1=.4; true values of network parameters: γ_0=0, γ_1=3, γ_2=-3; Z$\sim N(1,3)$; X$\sim$ AR(1), AR par=.8; rows stand for: (i): no correction applied , (ii) and (iv): linear and non-linear range as proxy, (iii) and (v): linear and non-linear standard deviation as proxy, (vi) and (vii): fixed effects and first-differencing estimators

4. Comparison across corrections, within (fe) and first-differencing (fd) estimators ($\rho = .7$)

Table 23: MSE and mean of parameters for models estimated using 2 and 4 lags of interaction variables

	MSE				Mean			
	ρ (2 lags)	β_1 (2 lags)	ρ (4 lags)	β_1 (4 lags)	ρ (2 lags)	β_1 (2 lags)	ρ (4 lags)	β_1 (4 lags)
no	0.27	0.00	0.12	0.00	1.01	0.40	0.99	0.40
range	0.01	0.00	0.00	0.00	0.72	0.40	0.72	0.40
sd	0.01	0.00	0.01	0.00	0.66	0.40	0.66	0.40
fe	0.12	0.00	0.04	0.00	0.73	0.40	0.73	0.40
fd	0.22	0.01	0.10	0.01	0.74	0.40	0.73	0.40

N=100, T=4, groups=10; true values of structural pars: ρ=.7 β_1=.4; true values of network parameters: γ_0=0, γ_1=1, γ_2=-1; Z$\sim N(1,3)$; X$\sim$ AR(1), AR par=.8; rows stand for: (i): no correction applied , (ii) and (iii): range and standard deviation as proxy, (iv) and (v): fixed effects and first-differencing estimators

Table 24: MSE and mean of parameters for models estimated using 2 and 4 lags of interaction variables (cont'd)

	MSE				Mean			
	ρ (2 lags)	β_1 (2 lags)	ρ (4 lags)	β_1 (4 lags)	ρ (2 lags)	β_1 (2 lags)	ρ (4 lags)	β_1 (4 lags)
no	0.20	0.00	0.13	0.00	1.00	0.40	0.99	0.40
range	0.01	0.00	0.01	0.00	0.72	0.40	0.72	0.40
sd	0.01	0.00	0.01	0.00	0.65	0.40	0.66	0.40
fe	0.15	0.00	0.05	0.00	0.72	0.40	0.72	0.40
fd	0.14	0.01	0.07	0.01	0.74	0.40	0.74	0.40

N=100, T=4, groups=10; true values of structural pars: ρ=.7 β_1=.4; true values of network parameters: γ_0=0, γ_1=1, γ_2=-2; Z$\sim N(1,3)$; X$\sim$ AR(1), AR par=.8; rows stand for: (i): no correction applied , (ii) and (iii): range and standard deviation as proxy, (iv) and (v): fixed effects and first-differencing estimators

Table 25: MSE and mean of parameters for models estimated using 2 and 4 lags of interaction variables (cont'd)

	MSE				Mean			
	ρ (2 lags)	β_1 (2 lags)	ρ (4 lags)	β_1 (4 lags)	ρ (2 lags)	β_1 (2 lags)	ρ (4 lags)	β_1 (4 lags)
no	0.21	0.00	0.12	0.00	0.99	0.40	0.99	0.40
range	0.01	0.00	0.01	0.00	0.71	0.40	0.72	0.40
sd	0.01	0.00	0.01	0.00	0.65	0.40	0.65	0.40
fe	0.13	0.00	0.05	0.00	0.72	0.40	0.72	0.40
fd	0.28	0.01	0.06	0.01	0.75	0.40	0.73	0.40

N=100, T=4, groups=10; true values of structural pars: ρ=.7 β_1=.4; true values of network parameters: γ_0=0, γ_1=1, γ_2=-3; Z$\sim N(1,3)$; X$\sim$ AR(1), AR par=.8; rows stand for: (i): no correction applied , (ii) and (iii): range and standard deviation as proxy, (iv) and (v): fixed effects and first-differencing estimators

Table 26: MSE and mean of parameters for models estimated using 2 and 4 lags of interaction variables (cont'd)

	MSE				Mean			
	ρ (2 lags)	β_1 (2 lags)	ρ (4 lags)	β_1 (4 lags)	ρ (2 lags)	β_1 (2 lags)	ρ (4 lags)	β_1 (4 lags)
no	0.67	0.00	0.14	0.00	0.90	0.40	0.91	0.40
range	0.00	0.00	0.00	0.00	0.70	0.40	0.70	0.40
sd	0.01	0.00	0.01	0.00	0.66	0.40	0.67	0.40
fe	0.07	0.00	0.02	0.00	0.71	0.40	0.71	0.40
fd	0.15	0.01	0.03	0.01	0.73	0.40	0.72	0.40

N=100, T=4, groups=10; true values of structural pars: ρ=.7 β_1=.4; true values of network parameters: γ_0=0, γ_1=2, γ_2=-1; Z∼ N(1,3); X∼ AR(1), AR par=.8; rows stand for: (i): no correction applied , (ii) and (iii): range and standard deviation as proxy, (iv) and (v): fixed effects and first-differencing estimators

Table 27: MSE and mean of parameters for models estimated using 2 and 4 lags of interaction variables (cont'd)

	MSE				Mean			
	ρ (2 lags)	β_1 (2 lags)	ρ (4 lags)	β_1 (4 lags)	ρ (2 lags)	β_1 (2 lags)	ρ (4 lags)	β_1 (4 lags)
no	0.64	0.01	0.15	0.00	0.91	0.40	0.92	0.40
range	0.00	0.00	0.00	0.00	0.70	0.40	0.70	0.40
sd	0.01	0.00	0.01	0.00	0.66	0.40	0.67	0.40
fe	0.07	0.00	0.02	0.00	0.71	0.40	0.71	0.40
fd	0.16	0.01	0.02	0.01	0.72	0.40	0.72	0.40

N=100, T=4, groups=10; true values of structural pars: ρ=.7 β_1=.4; true values of network parameters: γ_0=0, γ_1=2, γ_2=-2; Z∼ N(1,3); X∼ AR(1), AR par=.8; rows stand for: (i): no correction applied , (ii) and (iii): range and standard deviation as proxy, (iv) and (v): fixed effects and first-differencing estimators

Table 28: MSE and mean of parameters for models estimated using 2 and 4 lags of interaction variables (cont'd)

	MSE				Mean			
	ρ (2 lags)	β_1 (2 lags)	ρ (4 lags)	β_1 (4 lags)	ρ (2 lags)	β_1 (2 lags)	ρ (4 lags)	β_1 (4 lags)
no	0.44	0.00	0.13	0.00	0.92	0.40	0.92	0.40
range	0.00	0.00	0.00	0.00	0.69	0.40	0.70	0.40
sd	0.01	0.00	0.01	0.00	0.66	0.40	0.66	0.40
fe	0.25	0.00	0.02	0.00	0.71	0.40	0.71	0.40
fd	0.07	0.01	0.03	0.01	0.71	0.40	0.71	0.40

N=100, T=4, groups=10; true values of structural pars: ρ=.7 β_1=.4; true values of network parameters: γ_0=0, γ_1=2, γ_2=-3; Z∼ N(1,3); X∼ AR(1), AR par=.8; rows stand for: (i): no correction applied , (ii) and (iii): range and standard deviation as proxy, (iv) and (v): fixed effects and first-differencing estimators

Table 29: MSE and mean of parameters for models estimated using 2 and 4 lags of interaction variables (cont'd)

	MSE				Mean			
	ρ (2 lags)	β_1 (2 lags)	ρ (4 lags)	β_1 (4 lags)	ρ (2 lags)	β_1 (2 lags)	ρ (4 lags)	β_1 (4 lags)
no	0.26	0.00	0.11	0.00	0.87	0.40	0.87	0.40
range	0.00	0.00	0.00	0.00	0.69	0.40	0.69	0.40
sd	0.01	0.00	0.00	0.00	0.66	0.40	0.67	0.40
fe	0.04	0.00	0.01	0.00	0.70	0.40	0.70	0.40
fd	0.07	0.01	0.03	0.01	0.71	0.40	0.71	0.40

N=100, T=4, groups=10; true values of structural pars: ρ=.7 β_1=.4; true values of network parameters: γ_0=0, γ_1=3, γ_2=-1; Z∼ N(1,3); X∼ AR(1), AR par=.8; rows stand for: (i): no correction applied , (ii) and (iii): range and standard deviation as proxy, (iv) and (v): fixed effects and first-differencing estimators

Table 30: MSE and mean of parameters for models estimated using 2 and 4 lags of interaction variables (cont'd)

	MSE				Mean			
	ρ (2 lags)	β_1 (2 lags)	ρ (4 lags)	β_1 (4 lags)	ρ (2 lags)	β_1 (2 lags)	ρ (4 lags)	β_1 (4 lags)
no	0.30	0.00	0.14	0.00	0.88	0.40	0.88	0.40
range	0.00	0.00	0.00	0.00	0.68	0.40	0.69	0.40
sd	0.01	0.00	0.00	0.00	0.66	0.40	0.67	0.40
fe	0.03	0.00	0.01	0.00	0.70	0.40	0.71	0.40
fd	0.19	0.01	0.02	0.01	0.71	0.40	0.71	0.40

N=100, T=4, groups=10; true values of structural pars: ρ=.7 β_1=.4; true values of network parameters: γ_0=0, γ_1=3, γ_2=-2; Z∼ N(1,3); X∼ AR(1), AR par=.8; rows stand for: (i): no correction applied , (ii) and (iii): range and standard deviation as proxy, (iv) and (v): fixed effects and first-differencing estimators

Table 31: MSE and mean of parameters for models estimated using 2 and 4 lags of interaction variables (cont'd)

	MSE				Mean			
	ρ (2 lags)	β_1 (2 lags)	ρ (4 lags)	β_1 (4 lags)	ρ (2 lags)	β_1 (2 lags)	ρ (4 lags)	β_1 (4 lags)
no	0.52	0.00	0.11	0.00	0.88	0.40	0.87	0.40
range	0.00	0.00	0.00	0.00	0.68	0.40	0.69	0.40
sd	0.01	0.00	0.01	0.00	0.66	0.40	0.67	0.40
fe	0.07	0.00	0.01	0.00	0.71	0.40	0.71	0.40
fd	0.07	0.01	0.02	0.01	0.70	0.40	0.71	0.40

N=100, T=4, groups=10; true values of structural pars: ρ=.7 β_1=.4; true values of network parameters: γ_0=0, γ_1=3, γ_2=-3; Z∼ N(1,3); X∼ AR(1), AR par=.8; rows stand for: (i): no correction applied , (ii) and (iii): range and standard deviation as proxy, (iv) and (v): fixed effects and first-differencing estimators

5. Results using linear corrections for both endogenous network and covariate ($\rho = .3$)

Table 32: MSE and mean of parameters for models estimated using 2 and 4 lags of interaction variables

	MSE				Mean			
	ρ (2 lags)	β_1 (2 lags)	ρ (4 lags)	β_1 (4 lags)	ρ (2 lags)	β_1 (2 lags)	ρ (4 lags)	β_1 (4 lags)
no 1	0.25	0.01	0.22	0.01	0.71	0.45	0.72	0.45
no 2	0.22	0.00	0.21	0.00	0.74	0.39	0.75	0.39
range	0.01	0.00	0.01	0.00	0.33	0.40	0.33	0.40
sd	0.02	0.00	0.02	0.00	0.23	0.40	0.25	0.40

N=100, T=4, groups=10; true values of structural pars: ρ=.3 β_1=.4; true values of network parameters: γ_0=0, γ_1=1, γ_2=-1; Z$\sim$ N(1,3); X$\sim$ AR(1), AR par=.8; rows stand for: (i): no correction applied , (ii) time average of X used as correction for endogenous X (iii) and (iv): range and standard deviation as proxy for endogenous network and time average of X used as correction for endogenous X

Table 33: MSE and mean of parameters for models estimated using 2 and 4 lags of interaction variables (cont'd)

	MSE				Mean			
	ρ (2 lags)	β_1 (2 lags)	ρ (4 lags)	β_1 (4 lags)	ρ (2 lags)	β_1 (2 lags)	ρ (4 lags)	β_1 (4 lags)
no 1	0.26	0.01	0.21	0.01	0.70	0.46	0.72	0.45
no 2	0.21	0.00	0.21	0.00	0.74	0.39	0.75	0.39
range	0.01	0.00	0.01	0.00	0.33	0.40	0.33	0.40
sd	0.02	0.00	0.02	0.00	0.23	0.40	0.24	0.40

N=100, T=4, groups=10; true values of structural pars: ρ=.3 β_1=.4; true values of network parameters: γ_0=0, γ_1=1, γ_2=-2; Z$\sim$ N(1,3); X$\sim$ AR(1), AR par=.8; rows stand for: (i): no correction applied , (ii) time average of X used as correction for endogenous X (iii) and (iv): range and standard deviation as proxy for endogenous network and time average of X used as correction for endogenous X

Table 34: MSE and mean of parameters for models estimated using 2 and 4 lags of interaction variables (cont'd)

	MSE				Mean			
	ρ (2 lags)	β_1 (2 lags)	ρ (4 lags)	β_1 (4 lags)	ρ (2 lags)	β_1 (2 lags)	ρ (4 lags)	β_1 (4 lags)
no 1	0.27	0.01	0.21	0.01	0.70	0.46	0.72	0.45
no 2	0.21	0.00	0.21	0.00	0.73	0.39	0.74	0.39
range	0.01	0.00	0.01	0.00	0.32	0.40	0.32	0.40
sd	0.03	0.00	0.02	0.00	0.22	0.40	0.24	0.40

N=100, T=4, groups=10; true values of structural pars: ρ=.3 β_1=.4; true values of network parameters: γ_0=0, γ_1=1, γ_2=-3; Z$\sim$ N(1,3); X$\sim$ AR(1), AR par=.8; rows stand for: (i): no correction applied , (ii) time average of X used as correction for endogenous X (iii) and (iv): range and standard deviation as proxy for endogenous network and time average of X used as correction for endogenous X

Table 35: MSE and mean of parameters for models estimated using 2 and 4 lags of interaction variables (cont'd)

	MSE				Mean			
	ρ (2 lags)	β_1 (2 lags)	ρ (4 lags)	β_1 (4 lags)	ρ (2 lags)	β_1 (2 lags)	ρ (4 lags)	β_1 (4 lags)
no 1	0.32	0.01	0.20	0.01	0.62	0.47	0.62	0.47
no 2	0.18	0.01	0.15	0.01	0.65	0.39	0.65	0.39
range	0.01	0.00	0.01	0.00	0.30	0.40	0.30	0.40
sd	0.01	0.00	0.01	0.00	0.25	0.40	0.26	0.40

N=100, T=4, groups=10; true values of structural pars: ρ=.3 β_1=.4; true values of network parameters: γ_0=0, γ_1=2, γ_2=-1; Z$\sim$ N(1,3); X$\sim$ AR(1), AR par=.8; rows stand for: (i): no correction applied , (ii) time average of X used as correction for endogenous X (iii) and (iv): range and standard deviation as proxy for endogenous network and time average of X used as correction for endogenous X

Table 36: MSE and mean of parameters for models estimated using 2 and 4 lags of interaction variables (cont'd)

	MSE				Mean			
	ρ (2 lags)	β_1 (2 lags)	ρ (4 lags)	β_1 (4 lags)	ρ (2 lags)	β_1 (2 lags)	ρ (4 lags)	β_1 (4 lags)
no 1	0.31	0.01	0.19	0.01	0.62	0.47	0.62	0.47
no 2	0.18	0.01	0.15	0.01	0.65	0.39	0.65	0.39
range	0.01	0.00	0.01	0.00	0.29	0.40	0.30	0.40
sd	0.01	0.00	0.01	0.00	0.24	0.40	0.26	0.40

N=100, T=4, groups=10; true values of structural pars: ρ=.3 β_1=.4; true values of network parameters: γ_0=0, γ_1=2, γ_2=-2; Z$\sim$ N(1,3); X$\sim$ AR(1), AR par=.8; rows stand for: (i): no correction applied , (ii) time average of X used as correction for endogenous X (iii) and (iv): range and standard deviation as proxy for endogenous network and time average of X used as correction for endogenous X

Table 37: MSE and mean of parameters for models estimated using 2 and 4 lags of interaction variables (cont'd)

	MSE				Mean			
	ρ (2 lags)	β_1 (2 lags)	ρ (4 lags)	β_1 (4 lags)	ρ (2 lags)	β_1 (2 lags)	ρ (4 lags)	β_1 (4 lags)
no 1	0.30	0.01	0.18	0.01	0.62	0.47	0.62	0.47
no 2	0.17	0.01	0.15	0.01	0.65	0.39	0.65	0.39
range	0.01	0.00	0.01	0.00	0.29	0.40	0.30	0.40
sd	0.02	0.00	0.01	0.00	0.24	0.40	0.25	0.40

N=100, T=4, groups=10; true values of structural pars: ρ=.3 β_1=.4; true values of network parameters: γ_0=0, γ_1=2, γ_2=-3; Z$\sim$ N(1,3); X$\sim$ AR(1), AR par=.8; rows stand for: (i): no correction applied , (ii) time average of X used as correction for endogenous X (iii) and (iv): range and standard deviation as proxy for endogenous network and time average of X used as correction for endogenous X

Table 38: MSE and mean of parameters for models estimated using 2 and 4 lags of interaction variables (cont'd)

	MSE				Mean			
	ρ (2 lags)	β_1 (2 lags)	ρ (4 lags)	β_1 (4 lags)	ρ (2 lags)	β_1 (2 lags)	ρ (4 lags)	β_1 (4 lags)
no 1	0.36	0.01	0.19	0.01	0.59	0.47	0.59	0.47
no 2	0.17	0.01	0.13	0.01	0.62	0.39	0.60	0.39
range	0.01	0.00	0.01	0.00	0.28	0.40	0.28	0.40
sd	0.01	0.00	0.01	0.00	0.24	0.40	0.26	0.40

N=100, T=4, groups=10; true values of structural pars: ρ=.3 β_1=.4; true values of network parameters: γ_0=0, γ_1=3, γ_2=-1; Z∼ N(1,3); X∼ AR(1), AR par=.8; rows stand for: (i): no correction applied , (ii) time average of X used as correction for endogenous X (iii) and (iv): range and standard deviation as proxy for endogenous network and time average of X used as correction for endogenous X

Table 39: MSE and mean of parameters for models estimated using 2 and 4 lags of interaction variables (cont'd)

	MSE				Mean			
	ρ (2 lags)	β_1 (2 lags)	ρ (4 lags)	β_1 (4 lags)	ρ (2 lags)	β_1 (2 lags)	ρ (4 lags)	β_1 (4 lags)
no 1	0.39	0.01	0.18	0.01	0.58	0.47	0.59	0.47
no 2	0.17	0.01	0.13	0.01	0.62	0.39	0.60	0.39
range	0.01	0.00	0.01	0.00	0.28	0.40	0.28	0.40
sd	0.01	0.00	0.01	0.00	0.24	0.40	0.25	0.40

N=100, T=4, groups=10; true values of structural pars: ρ=.3 β_1=.4; true values of network parameters: γ_0=0, γ_1=3, γ_2=-2; Z∼ N(1,3); X∼ AR(1), AR par=.8; rows stand for: (i): no correction applied , (ii) time average of X used as correction for endogenous X (iii) and (iv): range and standard deviation as proxy for endogenous network and time average of X used as correction for endogenous X

Table 40: MSE and mean of parameters for models estimated using 2 and 4 lags of interaction variables (cont'd)

	MSE				Mean			
	ρ (2 lags)	β_1 (2 lags)	ρ (4 lags)	β_1 (4 lags)	ρ (2 lags)	β_1 (2 lags)	ρ (4 lags)	β_1 (4 lags)
no 1	0.31	0.01	0.18	0.01	0.59	0.47	0.58	0.47
no 2	0.17	0.01	0.13	0.01	0.62	0.39	0.60	0.39
range	0.01	0.00	0.01	0.00	0.27	0.40	0.28	0.40
sd	0.01	0.00	0.01	0.00	0.24	0.40	0.25	0.40

N=100, T=4, groups=10; true values of structural pars: ρ=.3 β_1=.4; true values of network parameters: γ_0=0, γ_1=3, γ_2=-3; Z∼ N(1,3); X∼ AR(1), AR par=.8; rows stand for: (i): no correction applied , (ii) time average of X used as correction for endogenous X (iii) and (iv): range and standard deviation as proxy for endogenous network and time average of X used as correction for endogenous X

Appendix 2: Further results

In this section we present some extensions with respect to the main body of results presented before. In particular, we let average group dimensionality to vary and we do so by i) increasing N while keeping fixed the number of groups and ii) keeping N fixed and reducing the number of groups. The first set of results is contained in Tables 40-48, while Tables 49-57 consider the second set of results. For both simulation exercises we compare the finite sample performances of our proposed estimators with fixed effect estimator.[23] The data generating process is the same described in the simulation section, but for cross-sectional dimension or the number of groups.

Both set of results confirm the goodness in a MSE sense of the estimators using network statistics to correct for the presence of unobserved heterogeneity. Either we consider the range or the standard deviation of the network, MSEs are extremely similar, no matter the lag-length used and the size of the network parameters γ_1 and γ_2. This holds even if empirical means for the two estimators may differ.

Results for the reference within-estimator are more complex, although the patterns are similar for results considering N=200 and those considering G=5. They clearly underperform for models estimated using two lags as instruments; when lag-length is four, their performance gets closer to "correction-based" estimators as the size of γ_1 increases.

1. Results for N=200

Table 41: MSE and mean of parameters for models estimated using 2 and 4 lags of interaction variables

	MSE				Mean			
	ρ (2 lags)	β_1 (2 lags)	ρ (4 lags)	β_1 (4 lags)	ρ (2 lags)	β_1 (2 lags)	ρ (4 lags)	β_1 (4 lags)
no	0.50	0.00	0.42	0.00	0.88	0.40	0.88	0.40
range	0.02	0.00	0.02	0.00	0.41	0.40	0.41	0.40
sd	0.02	0.00	0.01	0.00	0.27	0.40	0.27	0.40
fe	0.17	0.00	0.07	0.00	0.36	0.40	0.34	0.40

N=200, T=4, groups=10; true values of structural pars: ρ=.3 β_1=.4; true values of network parameters: γ_0=0, γ_1=1, γ_2=-1; Z$\sim$ N(1,3); X$\sim$ AR(1), AR par=.8; rows stand for: (i): no correction applied ,(ii) and (iii): range and standard deviation as proxy for endogenous network (iv): fixed effect

Table 42: MSE and mean of parameters for models estimated using 2 and 4 lags of interaction variables (cont'd)

	MSE				Mean			
	ρ (2 lags)	β_1 (2 lags)	ρ (4 lags)	β_1 (4 lags)	ρ (2 lags)	β_1 (2 lags)	ρ (4 lags)	β_1 (4 lags)
no	0.48	0.00	0.41	0.00	0.87	0.40	0.87	0.40
range	0.02	0.00	0.02	0.00	0.41	0.40	0.41	0.40
sd	0.02	0.00	0.02	0.00	0.27	0.40	0.27	0.40
fe	0.15	0.00	0.07	0.00	0.36	0.40	0.34	0.40

N=200, T=4, groups=10; true values of structural pars: ρ=.3 β_1=.4; true values of network parameters: γ_0=0, γ_1=1, γ_2=-2; Z$\sim$ N(1,3); X$\sim$ AR(1), AR par=.8; rows stand for: (i): no correction applied ,(ii) and (iii): range and standard deviation as proxy for endogenous network (iv): fixed effect

[23]We do not report results for the first-differencing estimator, based on the poorer records emerging from the main set of results.

Table 43: MSE and mean of parameters for models estimated using 2 and 4 lags of interaction variables (cont'd)

	MSE				Mean			
	ρ (2 lags)	β_1 (2 lags)	ρ (4 lags)	β_1 (4 lags)	ρ (2 lags)	β_1 (2 lags)	ρ (4 lags)	β_1 (4 lags)
no	0.51	0.00	0.40	0.00	0.85	0.40	0.87	0.40
range	0.02	0.00	0.02	0.00	0.41	0.40	0.41	0.40
sd	0.02	0.00	0.02	0.00	0.26	0.40	0.27	0.40
fe	0.15	0.00	0.07	0.00	0.36	0.40	0.35	0.40

N=200, T=4, groups=10; true values of structural pars: ρ=.3 β_1=.4; true values of network parameters: γ_0=0, γ_1=1, γ_2=-3; Z∼ N(1,3); X∼ AR(1), AR par=.8; rows stand for: (i): no correction applied ,(ii) and (iii): range and standard deviation as proxy for endogenous network (iv): fixed effect

Table 44: MSE and mean of parameters for models estimated using 2 and 4 lags of interaction variables (cont'd)

	MSE				Mean			
	ρ (2 lags)	β_1 (2 lags)	ρ (4 lags)	β_1 (4 lags)	ρ (2 lags)	β_1 (2 lags)	ρ (4 lags)	β_1 (4 lags)
no	0.70	0.00	0.34	0.00	0.76	0.40	0.72	0.40
range	0.01	0.00	0.01	0.00	0.34	0.40	0.33	0.40
sd	0.01	0.00	0.01	0.00	0.27	0.40	0.27	0.40
fe	0.16	0.00	0.03	0.00	0.31	0.40	0.30	0.40

N=200, T=4, groups=10; true values of structural pars: ρ=.3 β_1=.4; true values of network parameters: γ_0=0, γ_1=2, γ_2=-1; Z∼ N(1,3); X∼ AR(1), AR par=.8; rows stand for: (i): no correction applied ,(ii) and (iii): range and standard deviation as proxy for endogenous network (iv): fixed effect

Table 45: MSE and mean of parameters for models estimated using 2 and 4 lags of interaction variables (cont'd)

	MSE				Mean			
	ρ (2 lags)	β_1 (2 lags)	ρ (4 lags)	β_1 (4 lags)	ρ (2 lags)	β_1 (2 lags)	ρ (4 lags)	β_1 (4 lags)
no	0.75	0.00	0.34	0.00	0.73	0.40	0.73	0.40
range	0.01	0.00	0.01	0.00	0.34	0.40	0.33	0.40
sd	0.01	0.00	0.01	0.00	0.26	0.40	0.27	0.40
fe	0.20	0.00	0.03	0.00	0.30	0.40	0.30	0.40

N=200, T=4, groups=10; true values of structural pars: ρ=.3 β_1=.4; true values of network parameters: γ_0=0, γ_1=2, γ_2=-2; Z∼ N(1,3); X∼ AR(1), AR par=.8; rows stand for: (i): no correction applied ,(ii) and (iii): range and standard deviation as proxy for endogenous network (iv): fixed effect

Table 46: MSE and mean of parameters for models estimated using 2 and 4 lags of interaction variables (cont'd)

	MSE				Mean			
	ρ (2 lags)	β_1 (2 lags)	ρ (4 lags)	β_1 (4 lags)	ρ (2 lags)	β_1 (2 lags)	ρ (4 lags)	β_1 (4 lags)
no	0.68	0.00	0.33	0.00	0.74	0.40	0.73	0.40
range	0.01	0.00	0.01	0.00	0.34	0.40	0.34	0.40
sd	0.01	0.00	0.01	0.00	0.26	0.40	0.27	0.40
fe	0.50	0.00	0.03	0.00	0.28	0.40	0.31	0.40

N=200, T=4, groups=10; true values of structural pars: ρ=.3 β_1=.4; true values of network parameters: γ_0=0, γ_1=2, γ_2=-3; Z∼ N(1,3); X∼ AR(1), AR par=.8; rows stand for: (i): no correction applied ,(ii) and (iii): range and standard deviation as proxy for endogenous network (iv): fixed effect

Table 47: MSE and mean of parameters for models estimated using 2 and 4 lags of interaction variables (cont'd)

	MSE				Mean			
	ρ (2 lags)	β_1 (2 lags)	ρ (4 lags)	β_1 (4 lags)	ρ (2 lags)	β_1 (2 lags)	ρ (4 lags)	β_1 (4 lags)
no	1.19	0.00	0.30	0.00	0.68	0.40	0.62	0.40
range	0.01	0.00	0.00	0.00	0.29	0.40	0.29	0.40
sd	0.01	0.00	0.01	0.00	0.24	0.40	0.26	0.40
fe	8.16	0.00	0.02	0.00	0.41	0.40	0.30	0.40

N=200, T=4, groups=10; true values of structural pars: ρ=.3 β_1=.4; true values of network parameters: γ_0=0, γ_1=3, γ_2=-1; Z$\sim$ N(1,3); X$\sim$ AR(1), AR par=.8; rows stand for: (i): no correction applied ,(ii) and (iii): range and standard deviation as proxy for endogenous network (iv): fixed effect

Table 48: MSE and mean of parameters for models estimated using 2 and 4 lags of interaction variables (cont'd)

	MSE				Mean			
	ρ (2 lags)	β_1 (2 lags)	ρ (4 lags)	β_1 (4 lags)	ρ (2 lags)	β_1 (2 lags)	ρ (4 lags)	β_1 (4 lags)
no	1.45	0.00	0.30	0.00	0.65	0.40	0.62	0.40
range	0.01	0.00	0.01	0.00	0.29	0.40	0.30	0.40
sd	0.01	0.00	0.01	0.00	0.24	0.40	0.26	0.40
fe	0.20	0.00	0.02	0.00	0.32	0.40	0.31	0.40

N=200, T=4, groups=10; true values of structural pars: ρ=.3 β_1=.4; true values of network parameters: γ_0=0, γ_1=3, γ_2=-2; Z$\sim$ N(1,3); X$\sim$ AR(1), AR par=.8; rows stand for: (i): no correction applied ,(ii) and (iii): range and standard deviation as proxy for endogenous network (iv): fixed effect

Table 49: MSE and mean of parameters for models estimated using 2 and 4 lags of interaction variables (cont'd)

	MSE				Mean			
	ρ (2 lags)	β_1 (2 lags)	ρ (4 lags)	β_1 (4 lags)	ρ (2 lags)	β_1 (2 lags)	ρ (4 lags)	β_1 (4 lags)
no	1.10	0.00	0.30	0.00	0.61	0.40	0.63	0.40
range	0.01	0.00	0.01	0.00	0.30	0.40	0.30	0.40
sd	0.01	0.00	0.01	0.00	0.24	0.40	0.25	0.40
fe	0.08	0.00	0.02	0.00	0.31	0.40	0.31	0.40

N=200, T=4, groups=10; true values of structural pars: ρ=.3 β_1=.4; true values of network parameters: γ_0=0, γ_1=3, γ_2=-3; Z$\sim$ N(1,3); X$\sim$ AR(1), AR par=.8; rows stand for: (i): no correction applied ,(ii) and (iii): range and standard deviation as proxy for endogenous network (iv): fixed effect

5. Results for G=5

Table 50: MSE and mean of parameters for models estimated using 2 and 4 lags of interaction variables

	MSE				Mean			
	ρ (2 lags)	β_1 (2 lags)	ρ (4 lags)	β_1 (4 lags)	ρ (2 lags)	β_1 (2 lags)	ρ (4 lags)	β_1 (4 lags)
no	0.55	0.00	0.39	0.00	0.83	0.40	0.83	0.40
range	0.02	0.00	0.02	0.00	0.34	0.40	0.33	0.40
sd	0.04	0.00	0.03	0.00	0.22	0.40	0.23	0.40
fe	0.22	0.00	0.12	0.00	0.34	0.40	0.33	0.40

N=100, T=4, groups=5; true values of structural pars: ρ=.3 β_1=.4; true values of network parameters: γ_0=0, γ_1=1, γ_2=-1; Z∼ N(1,3); X∼ AR(1), AR par=.8; rows stand for: (i): no correction applied ,(ii) and (iii): range and standard deviation as proxy for endogenous network (iv): fixed effect

Table 51: MSE and mean of parameters for models estimated using 2 and 4 lags of interaction variables (cont'd)

	MSE				Mean			
	ρ (2 lags)	β_1 (2 lags)	ρ (4 lags)	β_1 (4 lags)	ρ (2 lags)	β_1 (2 lags)	ρ (4 lags)	β_1 (4 lags)
no	1.32	0.00	0.38	0.00	0.80	0.40	0.83	0.40
range	0.02	0.00	0.02	0.00	0.34	0.40	0.34	0.40
sd	0.04	0.00	0.03	0.00	0.21	0.40	0.23	0.40
fe	0.21	0.00	0.11	0.00	0.33	0.40	0.34	0.40

N=100, T=4, groups=5; true values of structural pars: ρ=.3 β_1=.4; true values of network parameters: γ_0=0, γ_1=1, γ_2=-2; Z∼ N(1,3); X∼ AR(1), AR par=.8; rows stand for: (i): no correction applied ,(ii) and (iii): range and standard deviation as proxy for endogenous network (iv): fixed effect

Table 52: MSE and mean of parameters for models estimated using 2 and 4 lags of interaction variables (cont'd)

	MSE				Mean			
	ρ (2 lags)	β_1 (2 lags)	ρ (4 lags)	β_1 (4 lags)	ρ (2 lags)	β_1 (2 lags)	ρ (4 lags)	β_1 (4 lags)
no	0.56	0.00	0.37	0.00	0.81	0.40	0.83	0.40
range	0.02	0.00	0.02	0.00	0.34	0.40	0.34	0.40
sd	0.04	0.00	0.03	0.00	0.21	0.40	0.23	0.40
fe	0.28	0.01	0.10	0.00	0.35	0.40	0.35	0.40

N=100, T=4, groups=5; true values of structural pars: ρ=.3 β_1=.4; true values of network parameters: γ_0=0, γ_1=1, γ_2=-3; Z∼ N(1,3); X∼ AR(1), AR par=.8; rows stand for: (i): no correction applied ,(ii) and (iii): range and standard deviation as proxy for endogenous network (iv): fixed effect

Table 53: MSE and mean of parameters for models estimated using 2 and 4 lags of interaction variables (cont'd)

	MSE				Mean			
	ρ (2 lags)	β_1 (2 lags)	ρ (4 lags)	β_1 (4 lags)	ρ (2 lags)	β_1 (2 lags)	ρ (4 lags)	β_1 (4 lags)
no	0.61	0.00	0.32	0.00	0.67	0.40	0.64	0.40
range	0.01	0.00	0.01	0.00	0.29	0.40	0.29	0.40
sd	0.02	0.00	0.02	0.00	0.23	0.40	0.24	0.40
fe	0.40	0.00	0.06	0.00	0.31	0.40	0.30	0.40

N=100, T=4, groups=5; true values of structural pars: ρ=.3 β_1=.4; true values of network parameters: γ_0=0, γ_1=2, γ_2=-1; Z∼ N(1,3); X∼ AR(1), AR par=.8; rows stand for: (i): no correction applied ,(ii) and (iii): range and standard deviation as proxy for endogenous network (iv): fixed effect

Table 54: MSE and mean of parameters for models estimated using 2 and 4 lags of interaction variables (cont'd)

	MSE				Mean			
	ρ (2 lags)	β_1 (2 lags)	ρ (4 lags)	β_1 (4 lags)	ρ (2 lags)	β_1 (2 lags)	ρ (4 lags)	β_1 (4 lags)
no	0.85	0.00	0.31	0.00	0.66	0.40	0.64	0.40
range	0.01	0.00	0.01	0.00	0.29	0.40	0.29	0.40
sd	0.02	0.00	0.02	0.00	0.23	0.40	0.24	0.40
fe	0.19	0.00	0.06	0.00	0.29	0.40	0.30	0.40

N=100, T=4, groups=5; true values of structural pars: ρ=.3 β_1=.4; true values of network parameters: γ_0=0, γ_1=2, γ_2=-2; Z∼ N(1,3); X∼ AR(1), AR par=.8; rows stand for: (i): no correction applied ,(ii) and (iii): range and standard deviation as proxy for endogenous network (iv): fixed effect

Table 55: MSE and mean of parameters for models estimated using 2 and 4 lags of interaction variables (cont'd)

	MSE				Mean			
	ρ (2 lags)	β_1 (2 lags)	ρ (4 lags)	β_1 (4 lags)	ρ (2 lags)	β_1 (2 lags)	ρ (4 lags)	β_1 (4 lags)
no	1.07	0.00	0.30	0.00	0.65	0.40	0.65	0.40
range	0.01	0.00	0.01	0.00	0.29	0.40	0.30	0.40
sd	0.02	0.00	0.02	0.00	0.23	0.40	0.24	0.40
fe	0.15	0.00	0.04	0.00	0.30	0.40	0.30	0.40

N=100, T=4, groups=5; true values of structural pars: ρ=.3 β_1=.4; true values of network parameters: γ_0=0, γ_1=2, γ_2=-3; Z∼ N(1,3); X∼ AR(1), AR par=.8; rows stand for: (i): no correction applied ,(ii) and (iii): range and standard deviation as proxy for endogenous network (iv): fixed effect

Table 56: MSE and mean of parameters for models estimated using 2 and 4 lags of interaction variables (cont'd)

	MSE				Mean			
	ρ (2 lags)	β_1 (2 lags)	ρ (4 lags)	β_1 (4 lags)	ρ (2 lags)	β_1 (2 lags)	ρ (4 lags)	β_1 (4 lags)
no	0.79	0.00	0.29	0.00	0.62	0.40	0.58	0.40
range	0.01	0.00	0.01	0.00	0.27	0.40	0.27	0.40
sd	0.02	0.00	0.02	0.00	0.22	0.40	0.24	0.40
fe	0.09	0.00	0.04	0.00	0.30	0.40	0.29	0.40

N=100, T=4, groups=5; true values of structural pars: ρ=.3 β_1=.4; true values of network parameters: γ_0=0, γ_1=3, γ_2=-1; Z∼ N(1,3); X∼ AR(1), AR par=.8; rows stand for: (i): no correction applied ,(ii) and (iii): range and standard deviation as proxy for endogenous network (iv): fixed effect

Table 57: MSE and mean of parameters for models estimated using 2 and 4 lags of interaction variables (cont'd)

	MSE				Mean			
	ρ (2 lags)	β_1 (2 lags)	ρ (4 lags)	β_1 (4 lags)	ρ (2 lags)	β_1 (2 lags)	ρ (4 lags)	β_1 (4 lags)
no	0.67	0.00	0.27	0.00	0.62	0.40	0.58	0.40
range	0.01	0.00	0.01	0.00	0.27	0.40	0.27	0.40
sd	0.02	0.00	0.02	0.00	0.22	0.40	0.24	0.40
fe	0.09	0.00	0.03	0.00	0.29	0.40	0.30	0.40

N=100, T=4, groups=5; true values of structural pars: ρ=.3 β_1=.4; true values of network parameters: γ_0=0, γ_1=3, γ_2=-2; Z$\sim$ N(1,3); X$\sim$ AR(1), AR par=.8; rows stand for: (i): no correction applied ,(ii) and (iii): range and standard deviation as proxy for endogenous network (iv): fixed effect

Table 58: MSE and mean of parameters for models estimated using 2 and 4 lags of interaction variables (cont'd)

	MSE				Mean			
	ρ (2 lags)	β_1 (2 lags)	ρ (4 lags)	β_1 (4 lags)	ρ (2 lags)	β_1 (2 lags)	ρ (4 lags)	β_1 (4 lags)
no	1.19	0.00	0.26	0.00	0.64	0.40	0.58	0.40
range	0.01	0.00	0.01	0.00	0.27	0.40	0.28	0.40
sd	0.02	0.00	0.01	0.00	0.22	0.40	0.24	0.40
fe	0.09	0.00	0.04	0.00	0.29	0.40	0.30	0.40

N=100, T=4, groups=5; true values of structural pars: ρ=.3 β_1=.4; true values of network parameters: γ_0=0, γ_1=3, γ_2=-3; Z$\sim$ N(1,3); X$\sim$ AR(1), AR par=.8; rows stand for: (i): no correction applied ,(ii) and (iii): range and standard deviation as proxy for endogenous network (iv): fixed effect

RECENT PUBLICATIONS BY *CEIS Tor Vergata*

Political Turnover and the Performance of Local Public Enterprises
Andrea De Meo and Lorenzo Ferrari
CEIS Research Paper, 438, August 2018

Home Sweet Home: the Effect of Sugar Protectionism on Emigration in Italy, 1876-1913
Carlo Ciccarelli, Alberto Dalmazzo and Daniela Vuri
CEIS Research Paper, 437, June 2018

Fiscal Compact and Debt Consolidation Dynamics
Luca Brugnolini and Luisa Corrado
CEIS Research Paper, 436, June 2018

Protection for Sale with Price Interactions and Incomplete Pass-Through
Barbara Annicchiarico and Enrico Marvasi
CEIS Research Paper, 435, June 2018

How Hard Is It to Maximise Profit? Evidence from a 19-th Century Italian State Monopoly
Carlo Ciccarelli, Gianni De Fraja and Silvia Tiezzi
CEIS Research Paper, 434, May 2018

Consumption Dynamics, Housing Collateral and Stabilisation Policies: AWay Forward for Policy Co-Ordination?
Jagjit S. Chadha, Germana Corrado and Luisa Corrado
CEIS Research Paper, 433, May 2018

Equilibrium Effort in Games with Homogeneous Production Functions and Homogeneous Valuation
Walter Ferrarese
CEIS Research Paper, 432, April 2018

Leniency, Asymmetric Punishment and Corruption: Evidence from China
Maria Perrotta Berlin, Bei Qin and Giancarlo Spagnolo
CEIS Research Paper, 431, April 2018

Detecting Co-Movements in Noncausal Time Series
Gianluca Cubadda, Alain Hecq and Sean Telg
CEIS Research Paper, 430, April 2018

We Can Be Heroes. Trust and Resilience in Corrupted Economic Environments
Leonardo Becchetti, Luca Corazzini and Vittorio Pelligra
CEIS Research Paper, 429, April 2018

DISTRIBUTION

Our publications are available online at www.ceistorvergata.it

MEDIA INQUIRIES AND INFORMATION

For media inquiries, please contact Barbara Piazzi at +39 06 72595652/01 or by e-mail at piazzi@ceis.uniroma2.it. Our web site, www.ceistorvergata.it, contains more information about Center's events, publications, and staff.

DEVELOPMENT AND SUPPORT

For information about contributing to CEIS Tor Vergata, please contact Carmen Tata at +39 06 72595615/01 or by e-mail at segr.ceis@economia.uniroma2.it