Dealing with Endogenous Shocks in Dynamic Friendship Network

Marchenko, Maria

DOI:
10.57938/977f182f-b49e-4182-9f09-ddb1615654cf

Published: 01/01/2019

Document Version
Publisher's PDF, also known as Version of record

Department of Economics
Working Paper No. 291

Dealing with Endogenous Shocks in Dynamic Friendship Network

Maria Marchenko

August 2019

WIRTSCHAFTS UNIVERSITÄT WIEN VIENNA UNIVERSITY OF ECONOMICS AND BUSINESS

Dealing with Endogenous Shocks in Dynamic Friendship Network

Maria Marchenko *

August 2, 2019

Abstract

Different types of shocks, or the treatment of one of the players in a specific network, may influence not only the future performance of themselves but also affect their network connections. It is crucial to explore the behaviour of the whole network in response to such an event. This paper focuses on the cases of endogenously formed shock. The logic used in the peer effect literature is adopted to develop the dynamic model and accounts for the endogeneity of the shock. The model allows us to predict the endogenous part of the shock and use the remaining unexpected component to estimate the effect of the shock on the changes in the performance of network connections. The identification conditions for effect are derived, and the consistent estimation procedure is proposed.

*JEL classification codes: C21, C49, I21
Please check the latest version at sites.google.com/site/mariavmarchenko/jmp.pdf
WU Vienna, Austria and National Research University Higher School of Economics, Moscow, Russia. Email: maria.marchenko@wu.ac.at.
This research has benefited from discussions with Andreas Dzemski, Maria Yudkevich, Alex Teytelboym, Andrea Weber, Enno Mammen, Jan Nimczik, Stephen Kastoryano, seminar participants in Mannheim and participants of RSSIA 2015.
The financial support from the Government of the Russian Federation within the framework of the Basic Research Program at the National Research University Higher School of Economics and within the framework of the implementation of the 5-100 Programme Roadmap of the National Research University Higher School of Economics is acknowledged.

1 Introduction

Different events happening to a network member influence not only their future behaviour and performance but also potentially affect the future outcomes of their peers. Network links, therefore, serve as a channel to transfer the shock from one person to the others. Such shocking events may be both completely unexpected or determined by the behaviour of the members of the network. The latter is of interest in this paper. Endogenous nature of such shocking events means that some of the characteristics that determine the performance of the network members will also influence the probability of the shock. Thus, it is impossible to identify the direct effect of the endogenous shock, if it is treated in the analysis in a similar way to the exogenous shock.

There are some examples in development literature, discussing the effect of the shocking event on the network performance, but the shock in such cases is exogenous individual treatment. For instance, Comola and Prina (2014) study randomized access to savings accounts in Nepal, taking into account the changes in the network structure and propose the dynamic peer effects model accounting for the network changes due to the intervention. This approach is only valid for exogenous shocks or interventions. In contrast, I propose the model, that allows for the shock to be endogenous. I am relying on the idea that the effect of the shock is carried by the random process that assigns the shock to otherwise similar network members, and that a shock can to a certain degree be predicted using the standard peer effect approach. Other things equal, following the peer effect logic, two similar network members will have the same probability of experiencing the shock, so that the difference between presence and absence of it for two such members will be determined by a random component. I propose the analysis in two steps. In the first step, the shock is disentangled into two components: the predicted probability of the shock and the unexpected component of the shock. The latter determines the changes in connections' future behaviour induced by the shock, as discussed earlier. The unexpected component is then included in the second step as one of the determinants of the future outcome of the peers.

To the best of my knowledge, this project is the first to introduce the dynamic peer effect in a social network model allowing for the presence of endogenous shocks[1]. Moreover, I provide the identification results for this model and propose an estimation procedure. The identification and estimation of the first step are straightforward adjustments of the approach by Bramoullé et al. (2009) after Lee (2003). It requires the existence of intransitive triads in the network corresponding to the assumption of no correlated effects, i.e., friends of friends are not connected to the student. Hence, the friends of friends do not affect the peer directly, but via the common friend only. The assumption of no correlated effects can be relaxed, using the stricter identifying assumption, as will be discussed in the paper.

The identification of the second step is novel and demonstrates the necessity of the longitudinal network variation. The identification relies, first, on the ability to calculate the random component of the shock, i.e., on the assumption that the first step model is valid, and second, on the existence of some exogeneity in network variation. Changes in the network allow for separate identification of the influence of the old and new network on the outcome, creating variation in the network characteristics necessary for identifying the effect of the shock component. It is worth mentioning that some of the network changes may happen as a result of the shocking event. However, changes in the network must not be driven solely by the shock. Some link changes should be caused by a natural adjustment in the social environment. Should this assumption not hold, the new network structure becomes dependent on the shock, not allowing to identify the effect of it. However, in most of the networks, it is safe to assume that the links are also changing due to reasons other than shocking events.

The identification results also suggest the estimation approach, namely the use of the characteristics of second (or third in case of presence of correlated effects) level of the

[1]See, for example, a review of the recent econometric literature on networks in Paula (2015)

network as instrumental variables. I provide the proof for consistency of such procedure. Same instrumental variables are often used in peer effects literature. See, for example, Bramoullé et al. (2009), De Giorgi et al. (2010), Nicoletti and Rabe (2019).

The paper is organized as follows. Section 2 introduces models and identification strategies. Section 3 describes the estimation approach. Section 4 provides the proof of the results. Section 5 concludes.

2 Model

2.1 Naïve approach

I propose a two-step model that allows estimating the effect of an unexpected event happening to network connections. Although I do not conduct the pure peer effect estimation, I use the classical peer effect framework as a starting point.

A naïve way to write down the dynamic peer effect model without modelling the link formation is just using the corresponding peer effect model for each period. The outcome variable depends on one's own exogenous characteristics, as well as on the average outcome of the peer group and average exogenous characteristics of peers.

$$y_i^1 = \alpha_1 + \beta_1 \sum_{j \neq i} G_{ij}^1 y_j^1 + \gamma_1 X_i^1 + \delta_1 \sum_{j \neq i} G_{ij}^1 X_j^1 + \xi_i + \epsilon_i^1, \qquad \mathbb{E}[\epsilon_i^1 | X^1] = 0, \tag{1}$$

$$y_i^2 = \alpha_2 + \beta_2 \sum_{j \neq i} G_{ij}^2 y_j^2 + \gamma_2 X_i^2 + \delta_2 \sum_{j \neq i} G_{ij}^2 X_j^2 + \xi_i + \epsilon_i^2, \qquad \mathbb{E}[\epsilon_i^2 | X^2] = 0, \tag{2}$$

where y_i^1 and y_i^2 are outcome variables of a network member i in the first and the second periods correspondingly.

X_i is a vector of individual exogenous characteristics affecting the outcome variable.

G_{ij}^1 and G_{ij}^2 are two adjacency matrices for the first and the second periods correspondingly, weighted by the number of links, and their entries have the value of $1/n_i$ if the link from i to j exists. Note that these matrices are not necessarily symmetrical, since no assumption is made for the network to be directed or undirected.

ξ_i - individual unobserved fixed characteristics, which may influence performance and link formation. It consists of the common for individual's connections unobservable component and individual's own unobserved fixed characteristics.

The unobserved individual characteristics reflect homophily of the individuals, which may influence both link formation and network outcomes. In the case of group interactions group fixed effects are often introduced to capture correlated effects, whereas in the case of interactions in the relatively big network, such as students' network, network fixed effects make little sense. Local differences, proposed by Bramoullé et al. (2009), may be used to address the issue of correlated effects. The dynamic structure of the data allows addressing this issue differently. The model can be written in terms of differences, eliminating possible unobserved fixed effect component in the error term.

$$\Delta y_i = \Delta\alpha + \beta_2 \sum_{j\neq i} G_{ij}^2 y_j^2 - \beta_1 \sum_{j\neq i} G_{ij}^1 y_j^1 + \gamma_2 X_i^2 - \gamma_1 X_i^1 + \delta_2 \sum_{j\neq i} G_{ij}^2 X_j^2 - \delta_1 \sum_{j\neq i} G_{ij}^1 X_j^1 + \Delta\epsilon_i$$

Assumption A. The outcome variable of a single period can be estimated using the one-period model.

This additional assumption allows avoiding the autoregressive component in the second-period model. The outcome of the previous period, therefore, has no effect on the current period. It is a valid assumption, once the model, estimating the current outcome, includes observed characteristics as well as endogenous and exogenous peer effects, and controls for the unobserved fixed characteristics. These elements are sufficient to predict outcomes in a lot of settings, for example, in educational framework.

The proposed model system 1 and 2, and consequently, the model in differences, can be further modified in order to catch the desired effect of a shock. Following the naïve approach, similar to the model by Comola and Prina (2014), one can write the model as follows.

The equation for the first period should remain unchanged:

$$y_i^1 = \alpha_1 + \beta_1 \sum_{j \neq i} G_{ij}^1 y_j^1 + \gamma_1 X_i^1 + \delta_1 \sum_{j \neq i} G_{ij}^1 X_j^1 + \xi_i + \epsilon_i^1,$$

Whereas, the second-period model shall take into account the shock of the connection's shocking event. The straightforward way to do so is just to include the binary variable in the vector of controls. Let D_i be a dummy variable for having any links with a shock in the first period.

$$y_i^2 = \alpha_2 + \beta_2 \sum_{j \neq i} G_{ij}^2 y_j^2 + \tilde{\delta} D_i + \gamma_2 X_i^2 + \delta_2 \sum_{j \neq i} G_{ij}^2 X_j^2 + \xi_i + \epsilon_i^{22}$$

The system can then be re-written in differences, eliminating the possible individual fixed effect:

$$\Delta y_i = (\alpha_2 - \alpha_1) + \beta_2 \sum_{j \neq i} G_{ij}^2 y_j^2 - \beta_1 \sum_{j \neq i} G_{ij}^1 y_j^1 + \tilde{\gamma} D_i +$$
$$+ \gamma_2 X_i^2 - \gamma_1 X_i^1 + \delta_2 \sum_{j \neq i} G_{ij}^2 X_j^2 - \delta_1 \sum_{j \neq i} G_{ij}^1 X_j^1 + \epsilon_i^2 - \epsilon_i^1$$

However, this type of the regression can only be estimated consistently if the shock is exogenous, as in the case of randomized treatment. As discussed in the introduction, I, on the contrary, want to look at the situations where the shock can be to some extent explained by the observed component of the model, and therefore, is endogenous. I propose to use the peer effect model of the first period to capture the predictable component of the probability of the shock and use only the remained unpredicted part to estimate the effect of the shock on the future performance.

Comola and Prina (2014) also model the changes of the network as a response to the exogenous treatment. At the moment, I am not modelling the link formation. The

[2] In general the coefficients in the model with the shock are different from the baseline one-period models (1) and (2), but I left the same notations for simplicity

variation of the network links is assumed and is a crucial identifying assumption. Importantly, I assume, that a significant part of the changes in the structure of the networks is caused by the individual characteristics and preferences and not solely by the shocking event.

Assumption B. Changes of the network as a response to unexpected shock can happen, however, there are exogenous changes in the network.

This assumption can potentially neglect part of the effect of the shock, translated indirectly via the changes in the network structure. The results in the following setting, therefore, rely only on the direct effect of the shock.

2.2 Proposed model with no correlated effects

2.2.1 The model

Taking into account all above-mentioned argument I estimate the following model at the first step:

$$P(retake_i) = \alpha + \beta \sum_{j \neq i} G_{ij}^1 y_j^1 + \gamma X_i^1 + \delta \sum_{j \neq i} G_{ij}^1 X_j^1 + \xi_i + \nu_i, \qquad \mathbb{E}[\nu_i | X^1] = 0 \qquad (3)$$

In this specification, the error term consists of two parts: unobserved correlated effect, and conditionally independent noise. Dynamic peer effect model will eliminate the correlated effect component in the second step of the model, leading to the conditional independence of the error term. However, on the first step in general $\mathbb{E}[\xi_i + \nu_i | X^1] \neq 0$. I will discuss two cases: assuming no correlated effects and with correlated effect. The latter will be considered in the later subsections. For the former, 3 is transformed as follows :

$$P(retake_i) = \alpha + \beta \sum_{j \neq i} G_{ij}^1 y_j^1 + \gamma X_i^1 + \delta \sum_{j \neq i} G_{ij}^1 X_j^1 + \nu_i, \qquad \mathbb{E}[\nu_i | X^1] = 0 \qquad (3a)$$

I then take the residuals of the equation 3a, which is the part of the probability of the connections' shock not predicted by the model. I then construct the aggregate shock of the connections for i as the combination of the residuals for the network of i. The baseline specification uses the average of the residuals: $UR_i = \sum_{j \neq i} G^1_{ij} \hat{\nu}_j$. However, other approaches to define UR_i could also be considered. The identification results and estimation procedure are not affected by the definition UR_i.

Then I am using it as an unexpected shock to plug-in in the following equation:

$$\Delta y_i = (\alpha_2 - \alpha_1) + \beta_2 \sum_{j \neq i} G^2_{ij} y^2_j - \beta_1 \sum_{j \neq i} G^1_{ij} y^1_j + \tilde{\boldsymbol{\delta}} UR_i + \gamma_2 X^2_i - \gamma_1 X^1_i +$$

$$+ \delta_2 \sum_{j \neq i} G^2_{ij} X^2_j - \delta_1 \sum_{j \neq i} G^1_{ij} X^1_j + \Delta \epsilon_i \tag{4}$$

Since the model in differences eliminates possible individual fixed effect component in error term, I am able to make a stricter assumption on the error term: $\mathbb{E}[\Delta \epsilon_i] = 0$. This condition will be used to prove the identification later.

$\tilde{\boldsymbol{\delta}}$ is the desired effect. It captures the influence of the unpredicted component of the connections' shocks on the changes in one's own performance.

Model in differences, additional to the elimination of individual fixed effect, may give a better interpretation in some settings. It estimates the changes of own performance in response to the shock additional to the changes of performance in comparison to the network, obtained by the single-period model.

Note that the model allows for a general setting with coefficients for the endogenous peer effect and exogenous characteristics being different in two periods: β_2 and β_1 and δ_2 and δ_1. Moreover, this also allows to take into account the changes in the network, looking at the influence of two different peer groups in two periods.

2.2.2 Identifying assumptions

The identification results for the first step of the model adopt Bramoullé et al. (2009) approach, whereas the result, obtained for the second stage, is, to the best of my knowledge, a novel result for the literature.

Lemma 1 *Let* $\gamma_1^2 + \delta_1^2 \neq 0$ *and* $\beta_1 \neq 0$[3]. *If matrices* I, G^1, $(G^1)^2$ *are linearly independent, coefficients in 3a are identified.*

The proof of Lemma 1 is given in Section 4. This is exactly the condition obtained by (Bramoullé et al., 2009), and can be proven similarly. The identification of the coefficients on the first step, hence, allow using the obtained residuals for the further analysis. The identification is ensured by the existence of intransitive triads in the network, i.e. the existence of a set of three individuals i, j, k such that i is influenced by j, j is influenced by k, but i is not influenced by k. This is a valid assumption for most networks.

The characteristics of the friends' friends can, therefore, be used as instrumental variables for the friends' outcomes. Due to the existence of intransitive triads, these instrumental variables will not directly influence one's own outcome.

Lemma 2 *In the case of no correlated effects, if the assumptions of Lemma 1 hold, if* $\gamma_2^2 + \delta_2^2 \neq 0$ *and* $\beta_2 \neq 0$[4], *if matrices* I, G^2, $(G^2)^2$ *are linearly independent, and if* $G^1 \neq G^2$, *and the changes in the network are mostly exogenous, coefficients in 4 are identified.*

Identification of Step 2 relies heavily on the variation in the network structure. However, it is important that some changes in the network are exogenous. This assumption is quite reasonable, for example, for the friendship networks in educational or labour setting. Students are likely to learn more about their classmates with time.

[3]These are the coefficients from the baseline peer effect model 1.
[4]The coefficients from the baseline peer effect model 2

Once there are new links formed in the next period, the variation between new and old connections help to capture the effect of the changes in the outcome. The identifying assumptions also put the restriction on the network matrix of the second period, as in the first period: the network should include intransitive triads. The proof of Lemma 2 can also be found in section 4.

2.3 Model with correlated effects

2.3.1 The model

As was already mentioned, the correlated effect appears due to the similar individual characteristics within a group. The correlated effect is unlikely to be present in big networks, however, once the network may suggest existence of smaller groups or subnetworks in it, the correlated effects are more likely to be present.

To deal with it and eliminate unobserved variables, I propose taking the local differences, i.e. averaging the equation 3 over the connections of i and subtracting this average from 3 and assuming that ξ_i are the same for the members of one smaller network, and hence, it will vanish after taking the local differences:

$$P(retake_i) - \sum_{j\neq i} G^1_{ij} P(retake_j) = \beta \sum_{j\neq i} G^1_{ij}[y^1_j - \sum_{k\neq j} G^1_{jk} y^1_k] + \gamma[X^1_i - \sum_{j\neq i} G^1_{ij} X^1_j] +$$

$$+ \delta \sum_{j\neq i} G^1_{ij}[X^1_j - \sum_{j\neq k} G^1_{jk} X^1_k] + \eta_i, \qquad \eta_i = [\nu_i - \sum_{j\neq i} G^1_{ij}\nu_j],\ \mathbb{E}[\eta_i|X^1] = 0 \quad (5)$$

Similarly to the case without correlated effects, I construct the shock for i, taking the average of their networks residuals: $UR_i = \sum_{j\neq i} G^1_{ij}\hat{\eta}_j$. The second stage is then identical to the case with no correlated effects:

$$\Delta y_i = (\alpha_2 - \alpha_1) + \beta_2 \sum_{j\neq i} G^2_{ij} y^2_j - \beta_1 \sum_{j\neq i} G^1_{ij} y^1_j + \tilde{\boldsymbol{\delta}} UR_i + \gamma_2 X^2_i - \gamma_1 X^1_i +$$

$$+ \delta_2 \sum_{j \neq i} G^2_{ij} X^2_j - \delta_1 \sum_{j \neq i} G^1_{ij} X^1_j + \Delta \epsilon_i \tag{6}$$

Model in differences, additional to the elimination of individual fixed effect, also gets rid off the correlated effects, therefore, no local differences are needed for the second stage equation.

2.3.2 Identifying assumptions

The identification results for the first step of the model again adopt Bramoullé et al. (2009) approach, whereas the result, obtained for the second stage, is new.

Lemma 3 *Let* $\gamma_1^2 + \delta_1^2 \neq 0$ *and* $\beta^1 \neq 0$[5]. *If matrices* I, G^1, $(G^1)^2$, $(G^1)^3$ *are linearly independent, coefficients in 5 are identified.*

The proof is given in Section 4. This condition again follows the result of (Bramoullé et al., 2009) in the presence of correlated effects, and can be proven in the similar manner. The identification of model with correlated effects is ensured by the existence of distances between two students of length 3 and more, i.e. the existence of a set of at least 4 individuals i, j, k, m such that i is influenced by j, j is influenced by k, k is influenced by m, but i is not influenced by both m and k, and j is not influenced by m. This is a bit more demanding assumption than in the case of no correlated effects, but still valid for a lot of networks' types.

This assumption also naturally suggests the identifying instruments for the estimation strategy. In the case with correlated effects exogenous characteristics of the friends' friends will be used to control for the correlated effects. Therefore, the next friends' circle will be required as an instrument for the friends' outcome.

Lemma 4 *In the case of correlated effects, if the assumptions of Lemma 3 hold, if* $\gamma_2^2 + \delta_2^2 \neq 0$ *and* $\beta_2 \neq 0$[6], *if matrices* I, G^2, $(G^2)^2$, $(G^2)^3$ *are linearly independent, and*

[5]The coefficients from the baseline peer effect model 1
[6]The coefficients from the baseline peer effect model 2

if $G^1 \neq G^2$, with changes not driven by the shock only, coefficients in 6 are identified.

Identification of Step 2 again heavily relies on the variation in the network structure. Moreover, the restrictions are put on the friendship matrix of the second period, requiring the distances between two students of length 3 and more. The proof of Lemma 4 is presented in Section 4.

3 Estimation strategy

3.1 No correlated effects

I first discuss the model that does not take into account correlation effects: 3a and 4.

Step 1. I partially repeat Bramoullé et al. (2009) for the first step and use the adaptation of Generalized 2SLS strategy proposed by Kelejian and Prucha (1998) and refined by Lee (2003). As the identification result suggests, $((\boldsymbol{G^1})^2\boldsymbol{X}, (\boldsymbol{G^1})^3\boldsymbol{X}, \dots)$ can be used as valid instruments to obtain consistent estimators.

First, recall the peer effect model in reduced form, written in matrix notations, offered in Bramoullé et al. (2009):

$$\boldsymbol{y^1} = \alpha_1\boldsymbol{i} + \beta_1\boldsymbol{G^1}\boldsymbol{y^1} + \gamma_1\boldsymbol{X^1} + \delta_1\boldsymbol{G^1}\boldsymbol{X^1} + \nu^1, \qquad \mathbb{E}[\nu|\boldsymbol{X^1}] = 0,$$

which gives

$$\mathbb{E}[\boldsymbol{G^1}\boldsymbol{y^1}|\boldsymbol{X^1}] = (\boldsymbol{I} - \beta_1\boldsymbol{G^1})^{-1}\boldsymbol{G^1}\alpha_1 + (\boldsymbol{I} - \beta_1\boldsymbol{G^1})^{-1}\boldsymbol{G^1}(\gamma_1\boldsymbol{I} + \delta_1\boldsymbol{G^1})\boldsymbol{X^1}$$

Note that the fist step model can be written as follows:

$$\boldsymbol{PR} = \alpha + \beta\boldsymbol{G^1}\boldsymbol{Y^1} + \gamma\boldsymbol{X^1} + \delta\boldsymbol{G^1}\boldsymbol{X^1} + \nu, \qquad \mathbb{E}[\nu|\boldsymbol{X^1}] = 0 \tag{7}$$

I propose the following procedure that gives the consistent estimator of $\theta = (\alpha, \beta, \gamma, \delta)$:

First, compute the 2SLS estimator for $\theta^1 = (\alpha_1, \beta_1, \gamma_1, \delta_1)$ of the standard peer effects model, using the following vector of instruments $\boldsymbol{S} = [\boldsymbol{i}, \boldsymbol{X^1}, \boldsymbol{G^1 X^1}, \boldsymbol{(G^1)^2 X^1}]$, and with the vector of covariates $\boldsymbol{\tilde{X}^1} = [\boldsymbol{i}, \boldsymbol{X^1}, \boldsymbol{G^1 X^1}, \boldsymbol{G^1 y^1}]$.

$\hat{\theta}^1_{2SLS} = (\boldsymbol{\tilde{X}^1}^T \boldsymbol{P_S} \boldsymbol{\tilde{X}^1})^{-1} \boldsymbol{\tilde{X}^1}^T \boldsymbol{P_S y^1}$, where $\boldsymbol{P_S} = \boldsymbol{S}(\boldsymbol{S}^T \boldsymbol{S})^{-1} \boldsymbol{S}^T$ is a projection matrix.

Second, define $\boldsymbol{\hat{Z}} = Z(\hat{\theta}^1_{2SLS}) = [\boldsymbol{i}, \boldsymbol{X^1}, \boldsymbol{G^1 X^1}, \mathbb{E}[\boldsymbol{G^1 y^1(\hat{\theta}^1_{2SLS})} | \boldsymbol{X^1}]]$,

where $\mathbb{E}[\boldsymbol{G^1 y^1(\hat{\theta}^1_{2SLS})} | \boldsymbol{X^1}] = \boldsymbol{G^1}(\boldsymbol{I} - \hat{\beta}_{1,2SLS} \boldsymbol{G^1})^{-1} \hat{\alpha}_{1,2SLS} + \boldsymbol{G^1}(\boldsymbol{I} - \hat{\beta}_{1,2SLS} \boldsymbol{G^1})^{-1}$ $(\hat{\gamma}_{1,2SLS} \boldsymbol{I} + \hat{\delta}_{1,2SLS} \boldsymbol{G^1}) \boldsymbol{X^1}$

Finally, use $\boldsymbol{\hat{Z}}$ as a vector of instruments to estimate 3a. Note that the vector of covariates coincides with the one used at the first step: $\boldsymbol{\tilde{X}^1}$. Then the following consistent estimator is obtained: $\hat{\theta}_{Lee} = (\boldsymbol{\hat{Z}}^T \boldsymbol{\tilde{X}^1})^{-1} \boldsymbol{\hat{Z}}^T \boldsymbol{PR}$.

This procedure is a modification of a procedure proposed in Lee (2003), therefore, the consistency result is closely related to his Theorem 1:

Lemma 5 *Under regularity conditions defined in Section 4, the estimator $\hat{\theta}_{Lee}$ is consistent and has the following limiting distribution,*

$$\sqrt{n}(\hat{\theta}_{Lee} - \theta) \xrightarrow{D} \mathcal{N}(0, \Psi), \tag{8}$$

with $\Psi = \sigma^2_\nu (lim_{n \to \infty} \frac{1}{n} \boldsymbol{Z}^T \boldsymbol{Z})^{-1}$ *and*
$\boldsymbol{Z} = [\boldsymbol{i}, \boldsymbol{X^1}, \boldsymbol{G^1 X^1}, \boldsymbol{G}^1(\boldsymbol{I} - \beta_1 \boldsymbol{G^1})^{-1} \alpha_1 + (\boldsymbol{I} - \beta_1 \boldsymbol{G^1})^{-1} (\gamma_1 \boldsymbol{I} + \delta_1 \boldsymbol{G^1}) \boldsymbol{X^1}]$

Discussion and detailed proof of the consistency of such estimator are given in Section 4.

Step 2. I am approaching the estimation of the second step also adopting the 2SLS procedure discussed for the first step. First, the model 4 can be rewritten in the following way:

$$\Delta \boldsymbol{y} = (\alpha_2 - \alpha_1)\boldsymbol{i} + \beta_2 \boldsymbol{G}^2 \boldsymbol{y^2} - \beta_1 \boldsymbol{G^1 y^1} + \boldsymbol{\tilde{\delta} UR} + \gamma_2 \boldsymbol{X^2_{TV}} - \gamma_1 \boldsymbol{X^1_{TV}} + \delta_2 \boldsymbol{G^2 X^2} -$$

$$- \delta_1 \boldsymbol{G^1 X^1} + \Delta\epsilon, \qquad \text{with } \boldsymbol{UR} \text{ defined as discussed in Section 2.2} \tag{9}$$

By X^1_{TV}, and X^2_{TV} I denote the subset of covariates, which are time-variant to avoid singularity problem of estimation.

Then a vector of covariates is as follows: $\boldsymbol{\bar{X}} = [\boldsymbol{i}, \boldsymbol{X^2_{TV}}, \boldsymbol{X^1_{TV}}, \boldsymbol{G^2X^2}, \boldsymbol{G^1X^1}, \boldsymbol{UR}, \boldsymbol{G^1y^1}, \boldsymbol{G^2y^2}]$. Following the logic of the first step I use $\boldsymbol{(G^2)^2X^2}$ as an instrument for $\boldsymbol{G^2y^2}$. However, $\mathbb{E}[(\boldsymbol{G^1y^1})^T\Delta\epsilon] \neq 0$, hence the instrument for $\boldsymbol{G^1y^1}$ is required. I propose to use $\mathbb{E}[\boldsymbol{G^1y^1(\hat{\theta}^1_{2SLS})}|\boldsymbol{X^1}]$ as an instrument, as obtained on the step 1. It is obvious that such an instrument is a valid instrument since it is uncorrelated with the second step error term and is clearly correlated with the outcome variable. Then I define $\boldsymbol{M} = [\boldsymbol{i}, \boldsymbol{X^2_{TV}}, \boldsymbol{X^1_{TV}}, \boldsymbol{G^2X^2}, \boldsymbol{G^1X^1}, \boldsymbol{UR}, \mathbb{E}[\boldsymbol{G^1y^1(\hat{\theta}^1_{2SLS})}|\boldsymbol{X^1}], \boldsymbol{(G^2)^2X^2}]$ as a vector of instruments.

I modify 9, taking expectations given $\boldsymbol{X^2}$ and recalling $\mathbb{E}[\Delta\epsilon] = 0$:

$$(\boldsymbol{I} - \beta_2\boldsymbol{G^2})\mathbb{E}[\boldsymbol{y^2}|\boldsymbol{X^2}] = (\alpha_2 - \alpha_1)\boldsymbol{i} + (\boldsymbol{I} - \beta_1\boldsymbol{G^1})\boldsymbol{y^1} + \boldsymbol{\tilde{\delta}UR} + \gamma_2\boldsymbol{X^2_{TV}} - \gamma_1\boldsymbol{X^1_{TV}} + \\ + \delta_2\boldsymbol{G^2X^2} - \delta_1\boldsymbol{G^1X^1}$$

$$\mathbb{E}[\boldsymbol{y^2}|\boldsymbol{X^2}] = (\boldsymbol{I} - \beta_2\boldsymbol{G^2})^{-1}[(\alpha_2 - \alpha_1)\boldsymbol{i} + (\boldsymbol{I} - \beta_1\boldsymbol{G^1})\boldsymbol{y^1} + \boldsymbol{\tilde{\delta}UR} + \gamma_2\boldsymbol{X^2_{TV}} - \gamma_1\boldsymbol{X^1_{TV}} + \\ + \delta_2\boldsymbol{G^2X^2} - \delta_1\boldsymbol{G^1X^1}]$$

Let $\mathbb{E}[\boldsymbol{G^2y^2}(\phi)|\boldsymbol{X^2}, \boldsymbol{X^1}] = \boldsymbol{G^2}(\boldsymbol{I} - \beta_2\boldsymbol{G^2})^{-1}[(\alpha_2 - \alpha_1)\boldsymbol{i} + (\boldsymbol{I} - \beta_1\boldsymbol{G^1})\mathbb{E}[\boldsymbol{y^1}(\theta^1)|\boldsymbol{X^1}] + \boldsymbol{\tilde{\delta}UR} + \gamma_2\boldsymbol{X^2_{TV}} - \gamma_1\boldsymbol{X^1_{TV}} + \delta_2\boldsymbol{G^2X^2} - \delta_1\boldsymbol{G^1X^1}]$, where $\mathbb{E}[\boldsymbol{y^1}(\theta^1)|\boldsymbol{X^1}] = \boldsymbol{G^2}(\boldsymbol{I} - \beta_1\boldsymbol{G^1})^{-1}\alpha_1 + (\boldsymbol{I} - \beta_1\boldsymbol{G^1})^{-1}(\gamma_1\boldsymbol{I} + \delta_1\boldsymbol{G^1})\boldsymbol{X^1}$.

Then I also define the following vector $\bar{\boldsymbol{Z}} = [\boldsymbol{i}, \boldsymbol{X^2_{TV}}, \boldsymbol{X^1_{TV}}, \boldsymbol{G^2X^2}, \boldsymbol{G^1X^1}, \boldsymbol{UR}, \mathbb{E}[\boldsymbol{G^1y^1}(\theta^1)|\boldsymbol{X^1}], \mathbb{E}[\boldsymbol{G^2y^2}(\phi)|\boldsymbol{X^2}, \boldsymbol{X^1}]$

I propose the following estimation procedure:

First, compute the 2SLS estimator for $\phi = (\alpha_1, \alpha_2, \beta_1, \beta_2, \gamma_1, \gamma_2, \delta_1, \delta_2)$ of the 9, using a vector of instruments $\boldsymbol{M}$ and a vector of covariates $\bar{\boldsymbol{X}}^{\boldsymbol{1}}$, as defined above. $\hat{\phi}^1_{2SLS} = (\bar{\boldsymbol{X}}^T \boldsymbol{P_M} \bar{\boldsymbol{X}})^{-1} \bar{\boldsymbol{X}}^T \boldsymbol{P_M}(\boldsymbol{y^2} - \boldsymbol{y^1})$, where $\boldsymbol{P_M} = \boldsymbol{M}(\boldsymbol{M}^T\boldsymbol{M})^{-1}\boldsymbol{M}^T$ is a projection matrix.

Second, define $\hat{\bar{\boldsymbol{Z}}} = \bar{\boldsymbol{Z}}(\hat{\phi}_{2SLS}) = [\boldsymbol{i}, \boldsymbol{X^2_{TV}}, \boldsymbol{X^1_{TV}}, \boldsymbol{G^2X^2}, \boldsymbol{G^1X^1}, \boldsymbol{UR}, \mathbb{E}[\boldsymbol{G^1y^1}(\hat{\theta}^1_{2SLS})|\boldsymbol{X^1}]], \mathbb{E}[\boldsymbol{G^2y^2}(\hat{\phi}_{2SLS})|\boldsymbol{X^2}, \boldsymbol{X^1}]$,
where $\mathbb{E}[\boldsymbol{G^1y^1}(\hat{\theta}^1_{2SLS})|\boldsymbol{X^1}] = (\boldsymbol{I} - \hat{\beta}_{1,2SLS}\boldsymbol{G^1})^{-1}\hat{\alpha}_{1,2SLS} + (\boldsymbol{I} - \hat{\beta}_{1,2SLS}\boldsymbol{G^1})^{-1}(\hat{\gamma}_{1,2SLS}\boldsymbol{I} + \hat{\delta}_{1,2SLS}\boldsymbol{G^1})\boldsymbol{X^1}$, with $\hat{\theta}^1_{2SLS}$ obtained as the estimation of the first stage on the first step.
and $\mathbb{E}[\boldsymbol{G^2y^2}(\hat{\phi}_{2SLS})|\boldsymbol{X^2}, \boldsymbol{X^1}] = \boldsymbol{G^2}(\boldsymbol{I} - \hat{\beta}_{2,2SLS}\boldsymbol{G^2})^{-1}[(\hat{\alpha}_{2,2SLS} - \hat{\alpha}_{1,2SLS})\boldsymbol{i} + (\boldsymbol{I} - \hat{\beta}_{1,2SLS}\boldsymbol{G^1})\mathbb{E}[\boldsymbol{y^1}(\hat{\theta}^1_{2SLS})|\boldsymbol{X^1}] + \hat{\hat{\boldsymbol{\delta}}}_{2SLS}\boldsymbol{UR} + \hat{\gamma}_{2,2SLS}\boldsymbol{X^2_{TV}} - \hat{\gamma}_{1,2SLS}\boldsymbol{X^1_{TV}} + \hat{\delta}_{2,2SLS}\boldsymbol{G^2X^2} - \hat{\delta}_{1,2SLS}\boldsymbol{G^1X^1}]$

Finally, I use $\hat{\bar{\boldsymbol{Z}}}$ as a new vector of instrument to estimate 9. Then the following consistent estimator is obtained: $\hat{\phi}_{Lee} = (\hat{\bar{\boldsymbol{Z}}}^T \bar{\boldsymbol{X}})^{-1} \hat{\bar{\boldsymbol{Z}}}^T (\boldsymbol{y^2} - \boldsymbol{y^1})$.

The consistency of this estimator is less straightforward, but it holds under the regularity conditions. The proof of the following Lemma is provided in Section 4.

Lemma 6 *Under regularity conditions defined in Section 4, the estimator $\hat{\phi}_{Lee}$ is consistent and has the following limiting distribution,*

$$\sqrt{n}(\hat{\phi}_{Lee} - \phi) \xrightarrow{D} \mathcal{N}(0, \Phi),$$

with $\Phi = (\sigma^2_{\epsilon^1} + \sigma^2_{\epsilon^2})(lim_{n\to\infty} \frac{1}{n} \bar{\boldsymbol{Z}}^T \bar{\boldsymbol{Z}})^{-1}$

3.2 Correlated effects

If the correlated effects are assumed to be present in the model the fist step model can be written as follows in matrix notation:

$$(\boldsymbol{I}-\boldsymbol{G^1})\boldsymbol{PR} = \beta(\boldsymbol{I}-\boldsymbol{G^1})\boldsymbol{G^1}\boldsymbol{Y^1} + \gamma(\boldsymbol{I}-\boldsymbol{G^1})\boldsymbol{X^1} + \delta(\boldsymbol{I}-\boldsymbol{G^1})\boldsymbol{G^1}\boldsymbol{X^1} + \eta,$$
$$\eta = (\boldsymbol{I}-\boldsymbol{G^1})\nu,\ \mathbb{E}[\eta|\boldsymbol{X^1}] = 0$$

I then use the peer effect model in local differences proposed in Bramoullé et al. (2009):

$$(\boldsymbol{I}-\boldsymbol{G^1})\boldsymbol{y^1} = \beta_1(\boldsymbol{I}-\boldsymbol{G^1})\boldsymbol{G^1}\boldsymbol{y^1} + \gamma_1(\boldsymbol{I}-\boldsymbol{G^1})\boldsymbol{X^1} + \delta_1(\boldsymbol{I}-\boldsymbol{G^1})\boldsymbol{G^1}\boldsymbol{X^1} + (\boldsymbol{I}-\boldsymbol{G^1})\nu^1,$$
$$\mathbb{E}[\nu^1|\boldsymbol{X^1}] = 0,$$

which gives

$$\mathbb{E}[(\boldsymbol{I}-\boldsymbol{G^1})\boldsymbol{G^1}\boldsymbol{y^1}|\boldsymbol{X^1}] = (\boldsymbol{I}-\beta_1\boldsymbol{G^1})^{-1}(\boldsymbol{I}-\boldsymbol{G^1})\boldsymbol{G^1}(\gamma_1\boldsymbol{I}+\delta_1\boldsymbol{G^1})\boldsymbol{X^1}$$

The proposed estimation procedure, in this case, is close to the first step with no correlated effects. I redo all the steps with the following vectors of instruments and covariates: instruments $\boldsymbol{S} = [(\boldsymbol{I}-\boldsymbol{G^1})\boldsymbol{X^1}, (\boldsymbol{I}-\boldsymbol{G^1})\boldsymbol{G^1}\boldsymbol{X^1}, (\boldsymbol{I}-\boldsymbol{G^1})\boldsymbol{(G^1)^2}\boldsymbol{X^1}]$ and covariates $\boldsymbol{\tilde{X}^1} = [(\boldsymbol{I}-\boldsymbol{G^1})\boldsymbol{X^1}, (\boldsymbol{I}-\boldsymbol{G^1})\boldsymbol{G^1}\boldsymbol{X^1}, (\boldsymbol{I}-\boldsymbol{G^1})\boldsymbol{G^1}\boldsymbol{y^1}]$.

Then I find the 2SLS estimator on the first step and use it to get the new vector of instruments: $\boldsymbol{\hat{Z}} = Z(\hat{\theta}^1_{2SLS}) = [(\boldsymbol{I}-\boldsymbol{G^1})\boldsymbol{X^1}, (\boldsymbol{I}-\boldsymbol{G^1})\boldsymbol{G^1}\boldsymbol{X^1}, \mathbb{E}[(\boldsymbol{I}-\boldsymbol{G^1})\boldsymbol{G^1}\boldsymbol{y^1}\boldsymbol{(\hat{\theta}^1_{2SLS})}|\boldsymbol{X^1}]]$, where $\mathbb{E}[(\boldsymbol{I}-\boldsymbol{G^1})\boldsymbol{G^1}\boldsymbol{y^1}\boldsymbol{(\hat{\theta}^1_{2SLS})}|\boldsymbol{X^1}] = \boldsymbol{G^1}(\boldsymbol{I}-\hat{\beta}_{1,2SLS}\boldsymbol{G^1})^{-1}(\boldsymbol{I}-\boldsymbol{G^1})(\hat{\gamma}_{1,2SLS}\boldsymbol{I}+$ $+\hat{\delta}_{1,2SLS}\boldsymbol{G^1})\boldsymbol{X^1}$.

The consistent estimator can then be obtained as follows: $\hat{\theta}_{Lee} = (\boldsymbol{\hat{Z}}^T\boldsymbol{\tilde{X}^1})^{-1}\boldsymbol{\hat{Z}}^T\boldsymbol{PR}$. Note that the proof of consistency follows directly by combining the result of Lee (2003) and the proof of Lemma 5, which can be found in Section 4.

Step 2 also requires some adjustments in this case. Due to the presence of correlated effects, $\mathbb{E}[\boldsymbol{G^1 y^1 (\hat{\theta}^1_{2SLS})}|\boldsymbol{X^1}]$ is no longer observable since it includes the unobserved fixed effects correlated with covariates and cannot be used as an instrument. Hence, I need to modify both vectors of covariates and instruments in the following way: $\bar{\boldsymbol{X}} = [(\boldsymbol{I}-\boldsymbol{G^1})\boldsymbol{X^2_{TV}}, (\boldsymbol{I}-\boldsymbol{G^1})\boldsymbol{X^1_{TV}}, (\boldsymbol{I}-\boldsymbol{G^1})\boldsymbol{G^2 X^2}, (\boldsymbol{I}-\boldsymbol{G^1})\boldsymbol{G^1 X^1}, (\boldsymbol{I}-\boldsymbol{G^1})\boldsymbol{UR}, (\boldsymbol{I}-\boldsymbol{G^1})\boldsymbol{G^1 y^1}, (\boldsymbol{I}-\boldsymbol{G^1})\boldsymbol{G^2 y^2}]$ is a new vector of covariates. I then use $\boldsymbol{(I-G^1)(G^2)^2 X^2}$ as an instrument for $(\boldsymbol{I}-\boldsymbol{G^1})\boldsymbol{G^2 y^2}$. I propose to use $\mathbb{E}[\boldsymbol{(I-G^1)G^1 y^1(\hat{\theta}^1_{2SLS})}|\boldsymbol{X}]$ as an instrument for $(\boldsymbol{I}-\boldsymbol{G^1})\boldsymbol{G^1 y^1}$. This instrument is clearly a valid instrument since it is uncorrelated with the second step error term and is clearly correlated with the outcome variable. Then I define $\boldsymbol{M} = [(\boldsymbol{I}-\boldsymbol{G^1})\boldsymbol{X^2_{TV}}, (\boldsymbol{I}-\boldsymbol{G^1})\boldsymbol{X^1_{TV}}, (\boldsymbol{I}-\boldsymbol{G^1})\boldsymbol{G^2 X^2}, (\boldsymbol{I}-\boldsymbol{G^1})\boldsymbol{G^1 X^1}, (\boldsymbol{I}-\boldsymbol{G^1})\boldsymbol{UR}, \mathbb{E}[\boldsymbol{(I-G^1)G^1 y^1(\hat{\theta}^1_{2SLS})}|\boldsymbol{X}], \boldsymbol{(I-G^1)(G^2)^2 X^2}]$ as a vector of instruments.

Applying the same changes to all relevant vectors, I then fully repeat the estimation procedure of the case of no correlated effects, and obtain the consistent estimator. Consistency of the estimator is achieved by the argument similar to the one in Lemma 6, proof of which and more detailed discussion on estimation procedure can be found in Section 4.

4 Proofs

Regularity conditions (adaptation of Lee (2003)):

Assumption 1. The matrices $(I-\beta^1 G^1)$ and $(I-\beta_2 G^2)$ are nonsingular

Assumption 2. The row and column sums of the matrices G^1, G^2, $(I-\beta^1 G^1)^{-1}$ and $(I-\beta_2 G^2)^{-1}$ are uniformly bounded in absolute value.

Assumption 3. The elements of the matrices X^1 and X^2 are uniformly bounded in absolute value

Assumption 4. The error terms $\{\nu_i : 1 \leq i \leq n\}$ are identically distributed. Furthermore, they are distributed (jointly) independently with $\mathbb{E}[\nu_i \boldsymbol{X_i^1}] = 0$ and $\mathbb{E}[\nu_i^2] = \sigma_\nu < \infty$. Additionally, they are assumed to possess finite fourth moments. The error terms $\{\Delta\epsilon_i : 1 \leq i \leq n\}$ are identically distributed. Furthermore, they are distributed (jointly) independently with $\mathbb{E}[\Delta\epsilon_i] = 0$ and $\mathbb{E}[\Delta\epsilon_i^2] = \sigma_{\epsilon^1} + \sigma_{\epsilon^2} < \infty$. Additionally, they are assumed to possess finite fourth moments

Assumption 5. The limit $\boldsymbol{J} = \lim_{n\to\infty} \frac{1}{n}\boldsymbol{Z}^T\boldsymbol{Z}$ exists and is nonsingular.

Assumption 6. The limit $\bar{\boldsymbol{J}} = \lim_{n\to\infty} \frac{1}{n}\bar{\boldsymbol{Z}}^T\bar{\boldsymbol{Z}}$ exists and is nonsingular.

Assumption 7. **Step 1.** The initial estimator β^1_{2SLS} of β_1 is n^a-consistent for some $a > 0$. The initial estimators α^1_{2SLS}, γ^1_{2SLS} and δ^1_{2SLS} are consistent estimators of α^1, γ^1 and δ^1, respectively. **Step 2.** The initial estimators $\beta_{1,2SLS}$ and $\beta_{2,2SLS}$ of β_1 and β_2 are n^b-consistent for some $b > 0$. The initial estimators $\alpha_{1,2SLS}$, $\alpha_{2,2SLS}$, $\gamma_{1,2SLS}$, $\gamma_{2,2SLS}$, $\delta_{1,2SLS}$ and $\delta_{2,2SLS}$ are consistent estimators of α_1, α_2, γ_1, γ_2, δ_1 and δ_2, respectively.

Proof of Lemma 1.

The structural form equation:

$$P(retake_i) = \alpha^1 + \beta^1 \sum_{j\neq i} G^1_{ij} y^1_j + \gamma^1 X^1_i + \delta^1 \sum_{j\neq i} G^1_{ij} X^1_j + \nu_i, \qquad \mathbb{E}[\nu_i|X] = 0$$

can be rewritten in the reduced form in the following manner:

$$\boldsymbol{PR} = \alpha^1 \boldsymbol{i} + \beta^1 \boldsymbol{G^1 y^1} + \gamma^1 \boldsymbol{X^1} + \delta^1 \boldsymbol{G^1 X^1} + \nu, \qquad \mathbb{E}[\nu|\boldsymbol{X^1}] = 0$$

$$\boldsymbol{PR} = \alpha^1 \boldsymbol{i} + \beta^1 \boldsymbol{G^1 y^1} + (\gamma^1 \boldsymbol{I} + \delta^1 \boldsymbol{G^1})\boldsymbol{X^1} + \nu, \qquad \mathbb{E}[\nu|\boldsymbol{X^1}] = 0$$

Taking conditional expectations:

$$\mathbb{E}[\boldsymbol{PR}|\boldsymbol{X^1}] = \alpha^1 \boldsymbol{i} + \beta^1 \boldsymbol{G^1}\mathbb{E}[\boldsymbol{y^1}|\boldsymbol{X^1}] + (\gamma^1 \boldsymbol{I} + \delta^1 \boldsymbol{G^1})\boldsymbol{X^1}$$

Note that $\boldsymbol{y}$ can be expressed in terms of peer effect model as the one used for the probability of shocks:

$$y_i^1 = \alpha_0 + \beta_0 \sum_{j \neq i} G_{ij}^1 y_j^1 + \gamma_0 X_i^1 + \delta_0 \sum_{j \neq i} G_{ij}^1 X_j^1 + \xi_i, \qquad \mathbb{E}[\xi_i | X] = 0$$

with reduced form:

$$\boldsymbol{y^1} = \alpha_0 \boldsymbol{i} + \beta_0 \boldsymbol{G^1} \boldsymbol{y^1} + (\gamma_0 \boldsymbol{I} + \delta_0 \boldsymbol{G^1}) \boldsymbol{X^1} + \xi, \qquad \mathbb{E}[\xi | \boldsymbol{X}] = 0$$

Then following steps of Bramoullé et al. (2009):

$$\boldsymbol{y^1} = \alpha_0 (\boldsymbol{I} - \beta_0 \boldsymbol{G^1})^{-1} + (\boldsymbol{I} - \beta_0 \boldsymbol{G^1})^{-1} (\gamma_0 \boldsymbol{I} + \delta_0 \boldsymbol{G^1}) \boldsymbol{X^1} + (\boldsymbol{I} - \beta_0 \boldsymbol{G^1})^{-1} \xi, \qquad \mathbb{E}[\xi | \boldsymbol{X}] = 0$$

Using $(\boldsymbol{I} - \beta_0 \boldsymbol{G^1})^{-1} = \sum_{k=0}^{\infty} \beta_0^k \boldsymbol{(G^1)}^k$:

$$\boldsymbol{y^1} = \alpha_0 (\boldsymbol{I} - \beta_0 \boldsymbol{G^1})^{-1} + \gamma_0 \boldsymbol{X^1} + (\gamma_0 \beta_0 + \delta_0) \sum_{k=0}^{\infty} \beta_0^k \boldsymbol{(G^1)}^{k+1} \boldsymbol{X^1} + \sum_{k=0}^{\infty} \beta_0^k \boldsymbol{(G^1)}^k \xi$$

And the expected mean friends' groups' performance conditional on $\boldsymbol{X^1}$ can be written as:

$$\mathbb{E}[\boldsymbol{G^1} \boldsymbol{y^1} | \boldsymbol{X^1}] = \alpha_0 (\boldsymbol{I} - \beta_0 \boldsymbol{G^1})^{-1} + \gamma_0 \boldsymbol{G^1} \boldsymbol{X^1} + (\gamma_0 \beta_0 + \delta_0) \sum_{k=0}^{\infty} \beta_0^k \boldsymbol{(G^1)}^{k+2} \boldsymbol{X^1}$$

As was proven in Bramoullé et al. (2009), if $\gamma_0 \beta_0 + \delta_0 \neq 0$ and I, G^1 and $(G^1)^2$ are linearly independent, the social effects are identified. So this expression can be plugged-in into the reduced form of the equation for the probability of retake.

$$\mathbb{E}[\boldsymbol{PR}|\boldsymbol{X^1}] = \alpha^1 \boldsymbol{i} + \beta^1(\alpha_0(\boldsymbol{I} - \beta_0 \boldsymbol{G^1})^{-1} + \gamma_0 \boldsymbol{G^1} \boldsymbol{X^1} + (\gamma_0\beta_0 + \delta_0) \sum_{k=0}^{\infty} \beta_0^k \boldsymbol{(G^1)}^{k+2} \boldsymbol{X^1}) +$$
$$+(\gamma^1 \boldsymbol{I} + \delta^1 \boldsymbol{G^1})\boldsymbol{X^1} = (\alpha^1 \boldsymbol{I} + \beta^1 \alpha_0 (\boldsymbol{I} - \beta_0 \boldsymbol{G^1})^{-1})) +$$
$$+\beta^1(\gamma_0\beta_0 + \delta_0) \sum_{k=0}^{\infty} \beta_0^k \boldsymbol{(G^1)}^{k+2} \boldsymbol{X^1} + (\gamma^1 \boldsymbol{I} + (\beta^1\gamma_0 + \delta^1)\boldsymbol{G^1})\boldsymbol{X^1}$$

or

$$\mathbb{E}[\boldsymbol{PR}|\boldsymbol{X^1}] = \alpha^1 \boldsymbol{I} + \beta^1(\alpha_0(\boldsymbol{I} - \beta_0 \boldsymbol{G^1})^{-1} + \beta^1(\boldsymbol{I} - \beta_0 \boldsymbol{G^1})^{-1}(\gamma_0 \boldsymbol{I} + \delta_0 \boldsymbol{G^1})\boldsymbol{G^1}\boldsymbol{X^1} + (\gamma^1 \boldsymbol{I} + \delta^1 \boldsymbol{G^1})\boldsymbol{X^1}$$

Now consider two sets of structural parameters $(\alpha^1, \beta^1, \gamma^1, \delta^1)$ and $(\tilde{\alpha}^1, \tilde{\beta}^1, \tilde{\gamma}^1, \tilde{\delta}^1)$ leading to the same reduced form. It means that:

$$\alpha^1 \boldsymbol{I} + \beta^1 \alpha_0 (\boldsymbol{I} - \beta_0 \boldsymbol{G^1})^{-1} = \tilde{\alpha}^1 \boldsymbol{I} + \tilde{\beta}^1 \alpha_0 (\boldsymbol{I} - \beta_0 \boldsymbol{G^1})^{-1}$$

$$\alpha^1 \boldsymbol{I} - \alpha^1 \beta_0 \boldsymbol{G^1} + \beta^1 \alpha_0 \boldsymbol{I} = \tilde{\alpha}^1 \boldsymbol{I} - \tilde{\alpha}^1 \beta_0 \boldsymbol{G^1} + \tilde{\beta}^1 \alpha_0 \boldsymbol{I}$$

$$(\alpha^1 - \tilde{\alpha}^1)\boldsymbol{I} + (\beta^1 \alpha_0 - \tilde{\beta}^1 \alpha_0)\boldsymbol{I} - (\alpha^1 \beta_0 - \tilde{\alpha}^1 \beta_0)\boldsymbol{G^1} = 0$$

$$(\alpha^1 - \tilde{\alpha}^1 + (\beta^1 - \tilde{\beta}^1)\alpha_0)\boldsymbol{I} = (\alpha^1 - \tilde{\alpha}^1)\beta_0 \boldsymbol{G^1}$$

and:

$$\beta^1(\boldsymbol{I} - \beta_0 \boldsymbol{G^1})^{-1}(\gamma_0 \boldsymbol{I} + \delta_0 \boldsymbol{G^1})\boldsymbol{G^1} + (\gamma^1 \boldsymbol{I} + \delta^1 \boldsymbol{G^1}) = \tilde{\beta}^1(\boldsymbol{I} - \beta_0 \boldsymbol{G^1})^{-1}(\gamma_0 \boldsymbol{I} + \delta_0 \boldsymbol{G^1})\boldsymbol{G^1} + (\tilde{\gamma}^1 \boldsymbol{I} + \tilde{\delta}^1 \boldsymbol{G^1})$$

$$\beta^1(\gamma_0 \boldsymbol{I} + \delta_0 \boldsymbol{G^1})\boldsymbol{G^1} + (\boldsymbol{I} - \beta_0 \boldsymbol{G^1})(\gamma^1 \boldsymbol{I} + \delta^1 \boldsymbol{G^1}) = \tilde{\beta}^1(\gamma_0 \boldsymbol{I} + \delta_0 \boldsymbol{G^1})\boldsymbol{G^1} + (\boldsymbol{I} - \beta_0 \boldsymbol{G^1})(\tilde{\gamma}^1 \boldsymbol{I} + \tilde{\delta}^1 \boldsymbol{G^1})$$

$$\beta^1 \gamma_0 \boldsymbol{G^1} + \beta^1 \delta_0 (\boldsymbol{G^1})^2 + (\gamma^1 \boldsymbol{I} - (\beta_0 \gamma^1 - \delta^1)\boldsymbol{G^1}) - \beta_0 \delta^1 \boldsymbol{(G^1)}^2) = \tilde{\beta}^1 \gamma_0 \boldsymbol{G^1} + \tilde{\beta}^1 \delta_0 (\boldsymbol{G^1}^2) +$$
$$+(\tilde{\gamma}^1 \boldsymbol{I} - (\beta_0 \tilde{\gamma}^1 - \tilde{\delta}^1)\boldsymbol{G^1}) - \beta_0 \tilde{\delta}^1 \boldsymbol{(G^1)}^2)$$

$$\gamma^1 \boldsymbol{I} + (\beta^1\gamma_0 + \beta^1\delta_0 - \beta_0\gamma^1 + \delta^1)\boldsymbol{G^1} - \beta_0\delta^1(\boldsymbol{G^1})^2 = \tilde{\gamma}^1 \boldsymbol{I} + (\tilde{\beta}^1\gamma_0 + \tilde{\beta}^1\delta_0 - \beta_0\tilde{\gamma}^1 + \tilde{\delta}^1)\boldsymbol{G^1} - \beta_0\tilde{\delta}^1(\boldsymbol{G^1})^2$$

$$(\gamma^1 - \tilde{\gamma}^1)\boldsymbol{I} + ((\beta^1 - \tilde{\beta}^1)\gamma_0 + (\beta^1 - \tilde{\beta}^1)\delta_0 - \beta_0(\gamma^1 - \tilde{\gamma}^1) + \delta^1 - \tilde{\delta}^1)\boldsymbol{G^1} + \beta_0(\tilde{\delta}^1 - \delta^1)(\boldsymbol{G^1})^2 = 0$$

Now let I, G^1 and $(G^1)^2$ be linearly independent. Then the above equality holds only if all three coefficients are 0:

$$\gamma^1 - \tilde{\gamma}^1 = 0$$
$$(\beta^1 - \tilde{\beta}^1)\gamma_0 + (\beta^1 - \tilde{\beta}^1)\delta_0 - \beta_0(\gamma^1 - \tilde{\gamma}^1) + \delta^1 - \tilde{\delta}^1 = 0$$
$$\beta_0(\tilde{\delta}^1 - \delta^1) = 0$$

If $\beta_0 \neq 0$ and $\gamma_0^2 + \delta_0^2 \neq 0$, two sets of coefficients $(\alpha^1, \beta^1, \gamma^1, \delta^1)$ and $(\tilde{\alpha}^1, \tilde{\beta}^1, \tilde{\gamma}^1, \tilde{\delta}^1)$ are equivalent. Note that the restrictions on the coefficients of the peer effect model suggest that the model has an endogenous peer effect and the performance depends on own set of observed characteristics, or on peers observed characteristics, or on both. These requirements are natural for the peer effect model and therefore, the identification result is achieved. ■

Proof of Lemma 2. (Identification, Step 2, no correlated effects)

Recall the second step equation:

$$\Delta y_i = (\alpha_2 - \alpha_1) + \beta_2 \sum_{j \neq i} G_{ij}^2 y_j^2 - \beta_1 \sum_{j \neq i} G_{ij}^1 y_j^1 + \tilde{\boldsymbol{\delta}} UR_i + \gamma_2 X_i^2 - \gamma_1 X_i^1 +$$

$$+ \delta_2 \sum_{j \neq i} G_{ij}^2 X_j^2 - \delta_1 \sum_{j \neq i} G_{ij}^1 X_j^1 + \Delta\epsilon_i$$

It can be rewritten in the reduced form as following:

$$\Delta \boldsymbol{y} = (\alpha_2 - \alpha_1)\boldsymbol{i} + \beta_2 \boldsymbol{G}^2 \boldsymbol{y^2} - \beta_1 \boldsymbol{G^1} \boldsymbol{y^1} + \tilde{\boldsymbol{\delta}} \boldsymbol{UR} + \gamma_2 \boldsymbol{X_{TV}^2} - \gamma_1 \boldsymbol{X_{TV}^1} + \delta_2 \boldsymbol{G^2} \boldsymbol{X^2} -$$

$$-\delta_1 \boldsymbol{G^1 X^1} + \Delta\epsilon, \qquad \text{with } \boldsymbol{UR} \text{ defined as discussed in Section 2 and } \mathbb{E}[\Delta\epsilon] = 0$$

This can be further modified in the following manner:

$$\mathbb{E}[\Delta \boldsymbol{y}|\boldsymbol{X^2}] = (\alpha_2 - \alpha_1)\boldsymbol{i} + \beta_2 \boldsymbol{G^2}\mathbb{E}[\boldsymbol{y^2}|\boldsymbol{X^2}] - \beta_1 \boldsymbol{G^1}\mathbb{E}[\boldsymbol{y^1}|\boldsymbol{X^2}] + \tilde{\boldsymbol{\delta}}\mathbb{E}[\boldsymbol{UR}|\boldsymbol{X^2}] +$$

$$+\gamma_2 \boldsymbol{X^2_{TV}} - \gamma_1 \boldsymbol{X^1_{TV}} + \delta_2 \boldsymbol{G^2 X^2} - \delta_1 \boldsymbol{G^1 X^1}$$

with

$$\mathbb{E}[\boldsymbol{y^1}|\boldsymbol{X^2}] = (\boldsymbol{I} - \beta_{0,1}\boldsymbol{G^1})^{-1}\alpha_{0,1} + (\boldsymbol{I} - \beta_{0,1}\boldsymbol{G^1})^{-1}(\gamma_{0,1}\boldsymbol{I} + \delta_{0,1}\boldsymbol{G^1})\mathbb{E}[\boldsymbol{X^1}|\boldsymbol{X^2}] =$$

$$= (\boldsymbol{I} - \beta_{0,1}\boldsymbol{G^1})^{-1}\alpha_{0,1} + (\boldsymbol{I} - \beta_{0,1}\boldsymbol{G^1})^{-1}(\gamma_{0,1}\boldsymbol{I} + \delta_{0,1}\boldsymbol{G^1})\boldsymbol{X^1},$$

since $\boldsymbol{X^1}$ is already known by the time $\boldsymbol{X^2}$ is revealed, therefore, the latter cannot add any new information.

Also:

$$\mathbb{E}[\boldsymbol{y^2}|\boldsymbol{X^2}] = (\boldsymbol{I} - \beta_{0,2}\boldsymbol{G^2})^{-1}\alpha_{0,2} + (\boldsymbol{I} - \beta_{0,2}\boldsymbol{G^2})^{-1}(\gamma_{0,2}\boldsymbol{I} + \delta_{0,2}\boldsymbol{G^2})\boldsymbol{X^2}$$

Note that $\boldsymbol{UR}$ is also defined at the first period, hence, the new information in $\boldsymbol{X^2}$ will not anything new for the expected value of the $\boldsymbol{UR}$, hence $\mathbb{E}[\boldsymbol{UR}|\boldsymbol{X^2}] = \boldsymbol{UR}$.

Also notice than in principle coefficients in the model in differences $\alpha_1, \beta_1, \gamma_1, \delta_1$, $\alpha_2, \beta_2, \gamma_2, \delta_2$ can be different from the corresponding coefficients in the single period peer effect models $\alpha_{0,1}, \beta_{0,1}, \gamma_{0,1}, \delta_{0,1}, \alpha_{0,2}, \beta_{0,2}, \gamma_{0,2}, \delta_{0,2}$. This can be due to the unaccounted in single period model fixed effects that can be eliminated in the model in differences and due to the presence of the shock in the model, which can take some of the effect, that would be otherwise attributed towards endogenous or exogenous effect.

Then, letting $\alpha = \alpha_2 - \alpha_1$

$$\mathbb{E}[\Delta \boldsymbol{y}|\boldsymbol{X^2}] = \alpha \boldsymbol{i} + \beta_2 \boldsymbol{G^2}(\boldsymbol{I} - \beta_{0,2}\boldsymbol{G^2})^{-1}(\alpha_{0,2} + (\gamma_{0,2}\boldsymbol{I} + \delta_{0,2}\boldsymbol{G^2})\boldsymbol{X^2}) - \beta_1 \boldsymbol{G^1}(\boldsymbol{I} - \beta_{0,1}\boldsymbol{G^1})^{-1}(\alpha_{0,1} +$$

$$(\gamma_{0,1}\boldsymbol{I} + \delta_{0,1}\boldsymbol{G^1})\boldsymbol{X^1}) + \tilde{\boldsymbol{\delta}}\boldsymbol{UR} + \gamma_2 \boldsymbol{X^2_{TV}} - \gamma_1 \boldsymbol{X^1_{TV}} + \delta_2 \boldsymbol{G^2}\boldsymbol{X^2} - \delta_1 \boldsymbol{G^1}\boldsymbol{X^1}$$

First, if $\boldsymbol{G^1} = \boldsymbol{G^1}$, then δ_2 and δ_1 are identified only partially, for time-variant variables of $\boldsymbol{X^1}$ and $\boldsymbol{X^2}$ respectively. This assumption can be relaxed, if we let the coefficients of the single period coincide with the coefficients of the coefficients of the model in differences. Then, however, the following assumption need to be made $\tilde{\delta} = 0$, meaning that the shock has no effect on the outcome, which is generally not true. Hence, $\boldsymbol{G^1} = \boldsymbol{G^1}$ is one of the identifying assumptions for the second step model.

Next, I follow similar steps to the proof of Lemma 1. Consider two sets of the parameters leading to the same reduced form, $(\alpha_1, \beta_1, \gamma_1, \delta_1,\ \alpha_2, \beta_2, \gamma_2, \delta_2,\ \tilde{\delta})$ and $(\tilde{\alpha_1}, \tilde{\beta_1}, \tilde{\gamma_1}, \tilde{\delta_1},$ $\tilde{\alpha_2}, \tilde{\beta_2}, \tilde{\gamma_2}, \tilde{\delta_2},\ \tilde{\tilde{\delta}})$. I do not include the single-period parameters, since their identification is achieved separately, if $\boldsymbol{I}, \boldsymbol{G^1}, (\boldsymbol{G^1})^2$ are linearly independent and if $\boldsymbol{I}, \boldsymbol{G^2}, (\boldsymbol{G^2})^2$ are also linearly independent. Then:

$$\alpha \boldsymbol{I} + \beta_2 \boldsymbol{G^2}(\boldsymbol{I} - \beta_{0,2}\boldsymbol{G^2})^{-1}\alpha_{0,2} - \beta_1 \boldsymbol{G^1}(\boldsymbol{I} - \beta_{0,1}\boldsymbol{G^1})^{-1}\alpha_{0,1} =$$

$$= \tilde{\alpha} \boldsymbol{I} + \tilde{\beta_2} \boldsymbol{G^2}(\boldsymbol{I} - \beta_{0,2}\boldsymbol{G^2})^{-1}\alpha_{0,2} - \tilde{\beta_1} \boldsymbol{G^1}(\boldsymbol{I} - \beta_{0,1}\boldsymbol{G^1})^{-1}\alpha_{0,1}$$

$$\beta_2 \boldsymbol{G^2}(\boldsymbol{I} - \beta_{0,2}\boldsymbol{G^2})^{-1}(\gamma_{0,2}\boldsymbol{I} + \delta_{0,2}\boldsymbol{G^2}) + (\gamma_2 \boldsymbol{I} + \delta_2 \boldsymbol{G^2}) =$$

$$= \tilde{\beta_2} \boldsymbol{G^2}(\boldsymbol{I} - \beta_{0,2}\boldsymbol{G^2})^{-1}(\gamma_{0,2}\boldsymbol{I} + \delta_{0,2}\boldsymbol{G^2}) + (\tilde{\gamma_2} \boldsymbol{I} + \tilde{\delta_2} \boldsymbol{G^2})$$

$$\beta_1 \boldsymbol{G^1}(\boldsymbol{I} - \beta_{0,1}\boldsymbol{G^1})^{-1}(\gamma_{0,1}\boldsymbol{I} + \delta_{0,1}\boldsymbol{G^1}) + (\gamma_1 \boldsymbol{I} + \delta_1 \boldsymbol{G^1}) =$$

$$= \tilde{\beta_1} \boldsymbol{G^1}(\boldsymbol{I} - \beta_{0,1}\boldsymbol{G^1})^{-1}(\gamma_{0,1}\boldsymbol{I} + \delta_{0,1}\boldsymbol{G^1}) + (\tilde{\gamma_1} \boldsymbol{I} + \tilde{\delta_1} \boldsymbol{G^1})$$

$$\tilde{\delta} = \tilde{\tilde{\delta}}$$

Note, that I added time invariant own exogenous variables to the vectors $\boldsymbol{X^1_{TV}}$ and $\boldsymbol{X^2_{TV}}$. Since they are not in the model, zeros are assumed on the additional elements of γ_1 and γ_2.

The third equation can be further simplified as following:

$$\gamma_1\boldsymbol{I} + (\delta_1 - \gamma_1\beta_{0,1} - \beta_1\gamma_{0,1})\boldsymbol{G^1} + (\beta_1\delta_{0,1} - \delta_1\beta_{0,1})\boldsymbol{G^1})^2 =$$

$$= \tilde{\gamma_1}\boldsymbol{I} + (\tilde{\delta_1} - \tilde{\gamma_1}\beta_{0,1} - \tilde{\beta_1}\gamma_{0,1})\boldsymbol{G^1} + (\tilde{\beta_1}\delta_{0,1} - \tilde{\delta_1}\beta_{0,1})\boldsymbol{G^1})^2$$

Then, if $\boldsymbol{I}, \boldsymbol{G^1}, (\boldsymbol{G^1})^2$ are linearly independent, the coefficients in front of these three matrices are 0:

$$\begin{aligned}
&\gamma_1 - \tilde{\gamma_1} = 0\\
&\delta_1 - \gamma_1\beta_{0,1} - \beta_1\gamma_{0,1} = \tilde{\delta_1} - \tilde{\gamma_1}\beta_{0,1} - \tilde{\beta_1}\gamma_{0,1}, \text{ or}\\
&(\delta_1 - \tilde{\delta_1}) - (\gamma_1 - \tilde{\gamma_1})\beta_{0,1} + (\beta_1 - \tilde{\beta_1})\gamma_{0,1} = 0\\
&\beta_1\delta_{0,1} - \delta_1\beta_{0,1} = \tilde{\beta_1}\delta_{0,1} - \tilde{\delta_1}\beta_{0,1}, \text{ or}\\
&(\beta_1 - \tilde{\beta_1})\delta_{0,1} - (\delta_1 - \tilde{\delta_1})\beta_{0,1} = 0
\end{aligned}$$

Now, if $\beta_{0,1} \neq 0$ and $\gamma_{0,1}^2 + \delta_{0,1}^2 \neq 0$, the two sets of the coefficients, $(\gamma_1, \delta_1, \beta_1)$ and $(\tilde{\gamma_1}, \tilde{\delta_1}, \tilde{\beta_1})$, coincide.

Similar argument is valid for the coefficient in front of $\boldsymbol{X^2}$, hence $(\gamma_2, \delta_2, \beta_2)$ and $(\tilde{\gamma_2}, \tilde{\delta_2}, \tilde{\beta_2})$, also coincide, when $\boldsymbol{I}, \boldsymbol{G^2}, (\boldsymbol{G^2})^2$ are linearly independent and $\beta_{0,2} \neq 0$ and $\gamma_{0,2}^2 + \delta_{0,2}^2 \neq 0$.

The other two equalities lead then automatically to $\alpha = \tilde{\alpha}$ and $\tilde{\delta} = \tilde{\tilde{\delta}}$ without any additional assumptions. Hence, the identification is achieved under the conditions of linear independence of $\boldsymbol{I}, \boldsymbol{G^1}, (\boldsymbol{G^1})^2$ and $\boldsymbol{I}, \boldsymbol{G^2}, (\boldsymbol{G^2})^2$ and $\boldsymbol{G^1} \neq \boldsymbol{G^2}$ and mentioned assumptions on the coefficients. ■

Proof of Lemma 3.

The structural form equation:

$$P(retake_i) - \sum_{j\neq i} G^1_{ij} P(retake_j) = \beta \sum_{j\neq i} G^1_{ij}[y^1_j - \sum_{k\neq j} G^1_{jk} y^1_k] + \gamma[X^1_i - \sum_{j\neq i} G^1_{ij} X^1_j] +$$
$$+\delta \sum_{j\neq i} G^1_{ij}[X^1_j - \sum_{k\neq j} G^1_{jk} X^1_k] + [\eta_i - \sum_{j\neq i} G^1_{ij}\eta_j], \qquad \mathbb{E}[\eta_i|X^1] = 0$$

can be rewritten in the reduced form in the following manner:

$$(\boldsymbol{I}-\boldsymbol{G^1})\boldsymbol{PR} = \beta(\boldsymbol{I}-\boldsymbol{G^1})\boldsymbol{G^1}\boldsymbol{y^1} + \gamma(\boldsymbol{I}-\boldsymbol{G^1})\boldsymbol{X^1} + \delta(\boldsymbol{I}-\boldsymbol{G^1})\boldsymbol{G^1}\boldsymbol{X^1} + (\boldsymbol{I}-\boldsymbol{G^1})\eta,\ \mathbb{E}[\eta|\boldsymbol{X^1}] = 0$$

$$(\boldsymbol{I}-\boldsymbol{G^1})\boldsymbol{PR} = \beta(\boldsymbol{I}-\boldsymbol{G^1})\boldsymbol{G^1}\boldsymbol{y^1} + (\gamma\boldsymbol{I} + \delta\boldsymbol{G^1})(\boldsymbol{I}-\boldsymbol{G^1})\boldsymbol{X^1} + (\boldsymbol{I}-\boldsymbol{G^1})\eta, \qquad \mathbb{E}[\eta|\boldsymbol{X^1}] = 0$$

Taking conditional expectations:

$$\mathbb{E}[(\boldsymbol{I}-\boldsymbol{G^1})\boldsymbol{PR}|\boldsymbol{X^1}] = \beta(\boldsymbol{I}-\boldsymbol{G^1})\boldsymbol{G^1}\mathbb{E}[\boldsymbol{y^1}|\boldsymbol{X^1}] + (\gamma\boldsymbol{I} + \delta\boldsymbol{G^1})(\boldsymbol{I}-\boldsymbol{G^1})\boldsymbol{X^1}$$

Note that $\boldsymbol{y}$ can be expressed in terms of peer effect model as the one used for the probability of shocks:

$$y^1_i - \sum_{j\neq i} G^1_{ij} y^1_j = \beta_0 \sum_{j\neq i} G^1_{ij}[y^1_j - \sum_{k\neq j} G^1_{jk} y^1_k] + \gamma_0[X^1_i - \sum_{j\neq i} G^1_{ij} X^1_j] + \delta_0 \sum_{j\neq i} G^1_{ij}[X^1_j -$$
$$- \sum_{k\neq j} G^1_{jk} X^1_k] + [\xi_i - \sum_{j\neq i} G^1_{ij}\xi_j], \quad \mathbb{E}[\xi_i|X^1] = 0$$

with reduced form:

$$(\boldsymbol{I}-\boldsymbol{G^1})\boldsymbol{y^1} = \beta_0(\boldsymbol{I}-\boldsymbol{G^1})\boldsymbol{G^1}\boldsymbol{y^1} + (\gamma_0\boldsymbol{I} + \delta_0\boldsymbol{G^1})(\boldsymbol{I}-\boldsymbol{G^1})\boldsymbol{X^1} + (\boldsymbol{I}-\boldsymbol{G^1})\xi, \quad \mathbb{E}[\xi|\boldsymbol{X^1}] = 0$$

Then following steps of Bramoullé et al. (2009):

$$(\boldsymbol{I}-\boldsymbol{G^1})\boldsymbol{y^1} = (\boldsymbol{I}-\beta_0\boldsymbol{G^1})^{-1}(\gamma_0\boldsymbol{I}+\delta_0\boldsymbol{G^1})(\boldsymbol{I}-\boldsymbol{G^1})\boldsymbol{X^1}+(\boldsymbol{I}-\beta_0\boldsymbol{G^1})^{-1}(\boldsymbol{I}-\boldsymbol{G^1})\xi, \quad \mathbb{E}[\xi|\boldsymbol{X}] = 0$$

And:

$$\mathbb{E}[(\boldsymbol{I}-\boldsymbol{G^1})\boldsymbol{G^1}\boldsymbol{y^1}|\boldsymbol{X^1}] = (\boldsymbol{I}-\beta_0\boldsymbol{G^1})^{-1}(\gamma_0\boldsymbol{I}+\delta_0\boldsymbol{G^1})(\boldsymbol{I}-\boldsymbol{G^1})\boldsymbol{G^1}\boldsymbol{X^1}$$

As was proven in Bramoullé et al. (2009), if $\gamma_0\beta_0 + \delta_0 \neq 0$ and $I, G^1, (G^1)^2$ and $(G^1)^3$ are linearly independent, the social effects are identified. So this expression can be plugged-in into the reduced form of the equation for the probability of retake.

$$\mathbb{E}[(\boldsymbol{I}-\boldsymbol{G^1})\boldsymbol{PR}|\boldsymbol{X^1}] = \beta((\boldsymbol{I}-\beta_0\boldsymbol{G^1})^{-1}(\gamma_0\boldsymbol{I}+\delta_0\boldsymbol{G^1})(\boldsymbol{I}-\boldsymbol{G^1})\boldsymbol{G^1})\boldsymbol{X^1}+(\gamma\boldsymbol{I}+\delta\boldsymbol{G^1})(\boldsymbol{I}-\boldsymbol{G^1})\boldsymbol{X^1}$$

Now consider two sets of structural parameters (β, γ, δ) and $(\tilde{\beta}, \tilde{\gamma}, \tilde{\delta})$ leading to the same reduced form. It means that:

$$\begin{aligned}
&\beta(\boldsymbol{I}-\beta_0\boldsymbol{G^1})^{-1}(\gamma_0\boldsymbol{I}+\delta_0\boldsymbol{G^1})(\boldsymbol{I}-\boldsymbol{G^1})\boldsymbol{G^1} + (\gamma\boldsymbol{I}+\delta\boldsymbol{G^1})(\boldsymbol{I}-\boldsymbol{G^1}) = \\
&= \tilde{\beta}(\boldsymbol{I}-\beta_0\boldsymbol{G^1})^{-1}(\gamma_0\boldsymbol{I}+\delta_0\boldsymbol{G^1})(\boldsymbol{I}-\boldsymbol{G^1})\boldsymbol{G^1} + (\tilde{\gamma}\boldsymbol{I}+\tilde{\delta}\boldsymbol{G^1})(\boldsymbol{I}-\boldsymbol{G^1})
\end{aligned}$$

$$\begin{aligned}
&\beta(\gamma_0\boldsymbol{I}+\delta_0\boldsymbol{G^1})(\boldsymbol{I}-\boldsymbol{G^1})\boldsymbol{G^1} + (\boldsymbol{I}-\beta_0\boldsymbol{G^1})(\gamma\boldsymbol{I}+\delta\boldsymbol{G^1})(\boldsymbol{I}-\boldsymbol{G^1}) = \\
&= \tilde{\beta}(\gamma_0\boldsymbol{I}+\delta_0\boldsymbol{G^1})(\boldsymbol{I}-\boldsymbol{G^1})\boldsymbol{G^1} + (\boldsymbol{I}-\beta_0\boldsymbol{G^1})(\tilde{\gamma}\boldsymbol{I}+\tilde{\delta}\boldsymbol{G^1})(\boldsymbol{I}-\boldsymbol{G^1})
\end{aligned}$$

$$\begin{aligned}
&\beta\gamma_0\boldsymbol{G^1} + (\beta\delta_0 - \beta\gamma_0)(\boldsymbol{G^1})^2 - \beta\delta_0(\boldsymbol{G^1})^3 + (\gamma\boldsymbol{I} - (\beta_0\gamma - \delta + \gamma)\boldsymbol{G^1}) - \\
&\qquad - (\beta_0\delta - \gamma\beta_0 + \delta)(\boldsymbol{G^1})^2 + \beta_0\delta(\boldsymbol{G^1})^3) = \\
&= \tilde{\beta}\gamma_0\boldsymbol{G^1} + (\tilde{\beta}\delta_0 - \tilde{\beta}\gamma_0)(\boldsymbol{G^1})^2 - \tilde{\beta}\delta_0(\boldsymbol{G^1})^3 + (\tilde{\gamma}\boldsymbol{I} - (\beta_0\tilde{\gamma} - \tilde{\delta} + \tilde{\gamma})\boldsymbol{G^1}) - \\
&\qquad - (\beta_0\tilde{\delta} - \tilde{\gamma}\beta_0 + \tilde{\delta})(\boldsymbol{G^1})^2 + \beta_0\tilde{\delta}(\boldsymbol{G^1})^3)
\end{aligned}$$

$$\gamma \boldsymbol{I} + (\beta\gamma_0 - \beta_0\gamma + \delta - \gamma)\boldsymbol{G^1} + (\beta\delta_0 - \beta\gamma_0 - \beta_0\delta + \beta_0\gamma - \delta)(\boldsymbol{G^1})^2 + (\beta_0\delta - \beta\delta_0)(\boldsymbol{G^1})^3 =$$
$$= \tilde{\gamma}\boldsymbol{I} + (\tilde{\beta}\gamma_0 - \beta_0\tilde{\gamma} + \tilde{\delta} - \tilde{\gamma})\boldsymbol{G^1} + (\tilde{\beta}\delta_0 - \tilde{\beta}\gamma_0 - \beta_0\tilde{\delta} + \beta_0\tilde{\gamma} - \tilde{\delta})(\boldsymbol{G^1})^2 + (\beta_0\tilde{\delta} - \tilde{\beta}\delta_0)(\boldsymbol{G^1})^3$$

$$(\gamma^1 - \tilde{\gamma}^1)\boldsymbol{I} + ((\beta - \tilde{\beta})\gamma_0 - (\gamma - \tilde{\gamma})\beta_0 + (\delta - \tilde{\delta}) - (\gamma - \tilde{\gamma}))\boldsymbol{G^1} + ((\beta - \tilde{\beta})\delta_0 - (\beta - \tilde{\beta})\gamma_0 -$$
$$-(\delta - \tilde{\delta})\beta_0 + (\gamma - \tilde{\gamma})\beta_0 - (\delta - \tilde{\delta}))(\boldsymbol{G^1})^2 + ((\delta - \tilde{\delta})\beta_0 - (\beta - \tilde{\beta})\delta_0)(\boldsymbol{G^1})^3 = 0$$

Now let I, G^1, $(G^1)^2$ and $(G^1)^3$ be linearly independent. Then the above equality holds only if all three coefficients are 0:

$$\gamma - \tilde{\gamma} = 0$$
$$(\beta - \tilde{\beta})\gamma_0 - (\gamma - \tilde{\gamma})\beta_0 + (\delta - \tilde{\delta}) - (\gamma - \tilde{\gamma}) = 0$$
$$(\beta - \tilde{\beta})\delta_0 - (\beta - \tilde{\beta})\gamma_0 - (\delta - \tilde{\delta})\beta_0 + (\gamma - \tilde{\gamma})\beta_0 - (\delta - \tilde{\delta}) = 0$$
$$(\delta - \tilde{\delta})\beta_0 - (\beta - \tilde{\beta})\delta_0 = 0$$

If $\beta_0 \neq 0$ and $\gamma_0^2 + \delta_0^2 \neq 0$, two sets of coefficients (β, γ, δ) and $(\tilde{\beta}, \tilde{\gamma}, \tilde{\delta})$ are equivalent. Note that the restrictions on the coefficients of the peer effect model suggest that the model has an endogenous peer effect and the performance depends on own set of observed characteristics, or on peers observed characteristics, or on both. These requirements are natural for the peer effect model and therefore, the identification result is achieved. ■

Proof of Lemma 4. (Identification, Step 2, correlated effects)

The proof for Lemma 4 follows directly by applying similar arguments to the proofs of Lemma 2 and Lemma 3. Then, the identification is achieved under the conditions of linear independence of $\boldsymbol{I}, \boldsymbol{G^1}, (\boldsymbol{G^1})^2, (\boldsymbol{G^1})^3$ and $\boldsymbol{I}, \boldsymbol{G^2}, (\boldsymbol{G^2})^2, (\boldsymbol{G^2})^3$ and $\boldsymbol{G^1} \neq \boldsymbol{G^2}$ and the following assumptions on the coefficients: $\beta_{0,1} \neq 0$, $\gamma_{0,1}^2 + \delta_{0,1}^2 \neq 0$, $\beta_{0,2} \neq 0$ and

$\gamma_{0,2}^2 + \delta_{0,2}^2 \neq 0$. ■

Proof of Lemma 5. Consistency of $\hat{\theta}_{Lee}$ of Step 1

$$\sqrt{n}(\hat{\theta}_{Lee}-\theta) = (\frac{1}{n}\hat{\boldsymbol{Z}}^T\tilde{\boldsymbol{X}}^{\mathbf{1}})^{-1}\frac{1}{\sqrt{n}}\hat{\boldsymbol{Z}}^T\boldsymbol{PR}-\sqrt{n}\theta = (\frac{1}{n}\hat{\boldsymbol{Z}}^T\tilde{\boldsymbol{X}}^{\mathbf{1}})^{-1}(\frac{1}{\sqrt{n}}\hat{\boldsymbol{Z}}^T\boldsymbol{PR}-\frac{1}{\sqrt{n}}\hat{\boldsymbol{Z}}^T\tilde{\boldsymbol{X}}^{\mathbf{1}}\theta)$$

Then we can rewrite the last term:

$$\frac{1}{\sqrt{n}}\hat{\boldsymbol{Z}}^T\boldsymbol{PR}-\frac{1}{\sqrt{n}}\hat{\boldsymbol{Z}}^T\tilde{\boldsymbol{X}}^{\mathbf{1}}\theta = \frac{1}{\sqrt{n}}\hat{\boldsymbol{Z}}^T(\boldsymbol{PR}-\tilde{\boldsymbol{X}}^{\mathbf{1}}\theta) = \frac{1}{\sqrt{n}}\hat{\boldsymbol{Z}}^T(\alpha\boldsymbol{i}+\beta\boldsymbol{G}^1\boldsymbol{y}^1+(\gamma\boldsymbol{I}+\delta\boldsymbol{G}^{\mathbf{1}})\boldsymbol{X}^{\mathbf{1}}+\nu-$$

$$-(\alpha\boldsymbol{i}+(\gamma\boldsymbol{I}+\delta\boldsymbol{G}^{\mathbf{1}})\boldsymbol{X}^{\mathbf{1}}+\beta\boldsymbol{G}^1\boldsymbol{y}^1)) = \frac{1}{\sqrt{n}}\hat{\boldsymbol{Z}}^T\nu$$

Hence,

$$\sqrt{n}(\hat{\theta}_{Lee}-\theta) = (\frac{1}{n}\hat{\boldsymbol{Z}}^T\tilde{\boldsymbol{X}}^{\mathbf{1}})^{-1}\frac{1}{\sqrt{n}}\hat{\boldsymbol{Z}}^T\nu$$

Then the following two statements can be shown under the assumed regularity conditions and by direct application of Lemmas A.7, A.8 and A.9 in Lee (2003):

$$plim\frac{1}{n}\hat{\boldsymbol{Z}}^T\tilde{\boldsymbol{X}}^{\mathbf{1}} = plim\frac{1}{n}\boldsymbol{Z}^T\boldsymbol{Z} = \boldsymbol{J}$$
$$\frac{1}{\sqrt{n}}\hat{\boldsymbol{Z}}^T\nu \xrightarrow{D} \mathcal{N}(0,\sigma_\nu^2\boldsymbol{J})$$

which will yield the desired result. ■

Proof of Lemma 6. Consistency of $\hat{\phi}_{Lee}$ of Step 2

$$\sqrt{n}(\hat{\phi}_{Lee}-\phi) = (\frac{1}{n}\hat{\bar{\boldsymbol{Z}}}^T\bar{\boldsymbol{X}})^{-1}\frac{1}{\sqrt{n}}\hat{\bar{\boldsymbol{Z}}}^T(\boldsymbol{y^2}-\boldsymbol{y^1})-\sqrt{n}\phi = (\frac{1}{n}\hat{\bar{\boldsymbol{Z}}}^T\bar{\boldsymbol{X}})^{-1}(\frac{1}{\sqrt{n}}\hat{\bar{\boldsymbol{Z}}}^T(\boldsymbol{y^2}-\boldsymbol{y^1})-\frac{1}{\sqrt{n}}\hat{\bar{\boldsymbol{Z}}}^T\bar{\boldsymbol{X}}\phi)$$

Then we can rewrite the last term:

$$\frac{1}{\sqrt{n}}\hat{\bar{\boldsymbol{Z}}}^T(\boldsymbol{y^2}-\boldsymbol{y^1})-\frac{1}{\sqrt{n}}\hat{\bar{\boldsymbol{Z}}}^T\bar{\boldsymbol{X}}\phi=\frac{1}{\sqrt{n}}\hat{\bar{\boldsymbol{Z}}}^T(\boldsymbol{y^2}-\boldsymbol{y^1}-\bar{\boldsymbol{X}}\phi)=\frac{1}{\sqrt{n}}\hat{\bar{\boldsymbol{Z}}}^T((\alpha_2-\alpha_1)\boldsymbol{i}+\beta_2\boldsymbol{G}^2\boldsymbol{y^2}-$$

$$-\beta_1\boldsymbol{G^1}\boldsymbol{y^1}+\tilde{\boldsymbol{\delta}}\boldsymbol{UR}+\gamma_2\boldsymbol{X^2_{TV}}-\gamma_1\boldsymbol{X^1_{TV}}+\delta_2\boldsymbol{G^2}\boldsymbol{X^2}-\delta_1\boldsymbol{G^1}\boldsymbol{X^1}+\Delta\epsilon-((\alpha_2-\alpha_1)\boldsymbol{i}+\beta_2\boldsymbol{G}^2\boldsymbol{y^2}-\beta_1\boldsymbol{G^1}\boldsymbol{y^1}+$$

$$+\tilde{\boldsymbol{\delta}}\boldsymbol{UR}+\gamma_2\boldsymbol{X^2_{TV}}-\gamma_1\boldsymbol{X^1_{TV}}+\delta_2\boldsymbol{G^2}\boldsymbol{X^2}-\delta_1\boldsymbol{G^1}\boldsymbol{X^1})=\frac{1}{\sqrt{n}}\hat{\bar{\boldsymbol{Z}}}^T\Delta\epsilon$$

Hence,

$$\sqrt{n}(\hat{\phi}_{Lee}-\phi)=(\frac{1}{n}\hat{\bar{\boldsymbol{Z}}}^T\bar{\boldsymbol{X}})^{-1}\frac{1}{\sqrt{n}}\hat{\bar{\boldsymbol{Z}}}^T\Delta\epsilon$$

The following two statements have to hold to get the desired result:

$$plim\frac{1}{n}\hat{\bar{\boldsymbol{Z}}}^T\bar{\boldsymbol{X}}=plim\frac{1}{n}\bar{\boldsymbol{Z}}^T\bar{\boldsymbol{Z}}=\bar{\boldsymbol{J}}$$
$$\frac{1}{\sqrt{n}}\hat{\bar{\boldsymbol{Z}}}^T\Delta\epsilon\xrightarrow{D}\mathcal{N}(0,(\sigma^2_{\epsilon^1}+\sigma^2_{\epsilon^2})\bar{\boldsymbol{J}})$$

First, let's consider $\frac{1}{n}\hat{\bar{\boldsymbol{Z}}}^T\bar{\boldsymbol{X}}$. It is equivalent to $\frac{1}{n}[\boldsymbol{i},\boldsymbol{X^2_{TV}},\boldsymbol{X^1_{TV}},\boldsymbol{G^2X^2},\boldsymbol{G^1X^1},\boldsymbol{UR},$ $\mathbb{E}[\boldsymbol{G^1y^1}(\hat{\theta}^1_{2SLS})|\boldsymbol{X^1}],\mathbb{E}[\boldsymbol{G^2y^2}(\hat{\phi}_{2SLS})|\boldsymbol{X^2},\boldsymbol{X^1}]]^T[\boldsymbol{i},\boldsymbol{X^2_{TV}},\boldsymbol{X^1_{TV}},\boldsymbol{G^2X^2},\boldsymbol{G^1X^1},\boldsymbol{UR},\boldsymbol{G^1y^1},$ $\boldsymbol{G^2y^2}]$

First six rows do not consist any element of estimated vector of coefficients, and therefore, will not matter for the consistency argument.

Notice also that $\boldsymbol{G^1y^1}=\boldsymbol{G^1}(\boldsymbol{I}-\beta^1\boldsymbol{G^1})^{-1}\alpha^1+\boldsymbol{G^1}(\boldsymbol{I}-\beta^1\boldsymbol{G^1})^{-1}(\gamma^1\boldsymbol{I}+\delta^1\boldsymbol{G^1})\boldsymbol{X^1}+$ $\boldsymbol{G^1}(\boldsymbol{I}-\beta^1\boldsymbol{G^1})^{-1}\epsilon_1$ and $\boldsymbol{G^2y^2}=\boldsymbol{G^2}(\boldsymbol{I}-\beta_2\boldsymbol{G^2})^{-1}[(\alpha_2-\alpha_1)\boldsymbol{i}+(\boldsymbol{I}-\beta_1\boldsymbol{G^1})((\boldsymbol{I}-\beta^1\boldsymbol{G^1})^{-1}\alpha^1+$ $(\boldsymbol{I}-\beta^1\boldsymbol{G^1})^{-1}(\gamma^1\boldsymbol{I}+\delta^1\boldsymbol{G^1})\boldsymbol{X^1})+\tilde{\boldsymbol{\delta}}\boldsymbol{UR}+\gamma_2\boldsymbol{X^2_{TV}}-\gamma_1\boldsymbol{X^1_{TV}}+\delta_2\boldsymbol{G^2X^2}-\delta_1\boldsymbol{G^1X^1}]+$ $\boldsymbol{G^2}(\boldsymbol{I}-\beta_2\boldsymbol{G^2})^{-1}\Delta\epsilon$ can be both split into two part: with and without error term.

Define $\mathbb{E}[\boldsymbol{G^1y^1}]\equiv\boldsymbol{G^1y^1}-\boldsymbol{G^1}(\boldsymbol{I}-\beta^1\boldsymbol{G^1})^{-1}\epsilon_1$ and $\mathbb{E}[\boldsymbol{G^2y^2}]\equiv\boldsymbol{G^2y^2}-\boldsymbol{G^2}(\boldsymbol{I}-\beta_2\boldsymbol{G^2})^{-1}\Delta\epsilon$

Consider now row six: $\frac{1}{n}(\mathbb{E}[\boldsymbol{G^1y^1}(\hat{\theta}^1_{2SLS})|\boldsymbol{X^1}])^T[\boldsymbol{i},\boldsymbol{X^2_{TV}},\boldsymbol{X^1_{TV}},\boldsymbol{G^2X^2},\boldsymbol{G^1X^1},\boldsymbol{UR},$ $\mathbb{E}[\boldsymbol{G^1y^1}],\mathbb{E}[\boldsymbol{G^2y^2}]]+\frac{1}{n}(\mathbb{E}[\boldsymbol{G^1y^1}(\hat{\theta}^1_{2SLS})|\boldsymbol{X^1}])^T[0,0,0,0,0,0,\boldsymbol{G^1}(\boldsymbol{I}-\beta^1\boldsymbol{G^1})^{-1}\epsilon_1,\boldsymbol{G^2}(\boldsymbol{I}-$

$\beta_2 \boldsymbol{G^2})^{-1}\Delta\epsilon]$.

By the assumed uniform boundedness of X^1, X^2 in absolute values as well as by the uniform boundness of the row and column sums of the matrices G^1, G^2, $(I-\beta^1 G^1)^{-1}$ and $(I-\beta_2 G^2)^{-1}$, by $\mathbb{E}[\Delta\epsilon]=0$ and by Lemmas A.6, A.7 and A.8 in Lee (2003), it can be shown that this row will have a limit in probability, which equals to corresponding row of $\bar{\boldsymbol{J}}$.

Similar argument holds for the row seven: $\frac{1}{n}(\mathbb{E}[\boldsymbol{G^2 y^2}(\hat{\phi}_{2SLS})|\boldsymbol{X^1}, \boldsymbol{X^2}])^T[\boldsymbol{i}, \boldsymbol{X^2_{TV}}, \boldsymbol{X^1_{TV}}, \boldsymbol{G^2 X^2}, \boldsymbol{G^1 X^1}, \boldsymbol{UR}, \mathbb{E}[\boldsymbol{G^1 y^1}], \mathbb{E}[\boldsymbol{G^2 y^2}]] + \frac{1}{n}(\mathbb{E}[\boldsymbol{G^2 y^2}(\hat{\phi}_{2SLS})|\boldsymbol{X^1}, \boldsymbol{X^2}])^T[0,0,0,0,0,0, \boldsymbol{G^1}(\boldsymbol{I}-\beta^1 \boldsymbol{G^1})^{-1}\epsilon_1, \boldsymbol{G^2}(\boldsymbol{I}-\beta_2 \boldsymbol{G^2})^{-1}\Delta\epsilon]$. Therefore, the first statement is correct.

For the second statement consider $\frac{1}{\sqrt{n}}\hat{\bar{\boldsymbol{Z}}}^T\Delta\epsilon = [\boldsymbol{i}, \boldsymbol{X^2_{TV}}, \boldsymbol{X^1_{TV}}, \boldsymbol{G^2 X^2}, \boldsymbol{G^1 X^1}, \boldsymbol{UR}, \mathbb{E}[\boldsymbol{G^1 y^1}(\hat{\theta}^1_{2SLS})|\boldsymbol{X^1}], \mathbb{E}[\boldsymbol{G^2 y^2}(\hat{\phi}_{2SLS})|\boldsymbol{X^2}, \boldsymbol{X^1}]]^T\Delta\epsilon$.

None of the elements in $\hat{\bar{\boldsymbol{Z}}}$ consist $\Delta\epsilon$, therefore, since $\mathbb{E}[\Delta\epsilon]=0$, the expectation of the whole term gives 0, which concludes the consistency part of the proof.

Moreover, the variance can be written as $(\sigma_{\epsilon^1}+\sigma_{\epsilon^2})\mathbb{E}[\frac{1}{n}\hat{\bar{\boldsymbol{Z}}}^T\hat{\bar{\boldsymbol{Z}}}]$. By the same Lemmas as before, it can be shown that $plim\mathbb{E}[\frac{1}{n}\hat{\bar{\boldsymbol{Z}}}^T\hat{\bar{\boldsymbol{Z}}}] = plim\frac{1}{n}\bar{\boldsymbol{Z}}^T\bar{\boldsymbol{Z}} = \bar{\boldsymbol{J}}$, which concludes the proof of normality.■.

Discussion of 2.4.2, step 2.

I am approaching the estimation of the second step also adopting the 2SLS procedure discussed for the first step. First, the model (5) can be rewritten in the following way:

$$\Delta\boldsymbol{y} = (\alpha_2-\alpha_1)\boldsymbol{i} + \beta_2\boldsymbol{G}^2\boldsymbol{y^2} - \beta_1\boldsymbol{G^1 y^1} + \tilde{\boldsymbol{\delta}}\boldsymbol{UR} + \gamma_2\boldsymbol{X^2_{TV}} - \gamma_1\boldsymbol{X^1_{TV}} + \delta_2\boldsymbol{G^2 X^2} -$$

$$-\delta_1\boldsymbol{G^1 X^1} + \Delta\epsilon$$

Then:

$$(\boldsymbol{I}-\boldsymbol{G^1})\Delta \boldsymbol{y} = \beta_2(\boldsymbol{I}-\boldsymbol{G^1})\boldsymbol{G}^2\boldsymbol{y^2} - \beta_1(\boldsymbol{I}-\boldsymbol{G^1})\boldsymbol{G^1}\boldsymbol{y^1} + \tilde{\boldsymbol{\delta}}(\boldsymbol{I}-\boldsymbol{G^1})\boldsymbol{UR} + \gamma_2(\boldsymbol{I}-\boldsymbol{G^1})\boldsymbol{X^2_{TV}} -$$

$$-\gamma_1(\boldsymbol{I}-\boldsymbol{G^1})\boldsymbol{X^1_{TV}} + \delta_2(\boldsymbol{I}-\boldsymbol{G^1})\boldsymbol{G^2}\boldsymbol{X^2} - \delta_1(\boldsymbol{I}-\boldsymbol{G^1})\boldsymbol{G^1}\boldsymbol{X^1} + (\boldsymbol{I}-\boldsymbol{G^1})\Delta\epsilon \quad (10)$$

Recall: $\bar{\boldsymbol{X}} = [(\boldsymbol{I}-\boldsymbol{G^1})\boldsymbol{X^2_{TV}}, (\boldsymbol{I}-\boldsymbol{G^1})\boldsymbol{X^1_{TV}}, (\boldsymbol{I}-\boldsymbol{G^1})\boldsymbol{G^2}\boldsymbol{X^2}, (\boldsymbol{I}-\boldsymbol{G^1})\boldsymbol{G^1}\boldsymbol{X^1}, (\boldsymbol{I}-\boldsymbol{G^1})\boldsymbol{UR}, (\boldsymbol{I}-\boldsymbol{G^1})\boldsymbol{G^1}\boldsymbol{y^1}, (\boldsymbol{I}-\boldsymbol{G^1})\boldsymbol{G^2}\boldsymbol{y^2}]$.

And $\boldsymbol{M} = [(\boldsymbol{I}-\boldsymbol{G^1})\boldsymbol{X^2_{TV}}, (\boldsymbol{I}-\boldsymbol{G^1})\boldsymbol{X^1_{TV}}, (\boldsymbol{I}-\boldsymbol{G^1})\boldsymbol{G^2}\boldsymbol{X^2}, (\boldsymbol{I}-\boldsymbol{G^1})\boldsymbol{G^1}\boldsymbol{X^1}, (\boldsymbol{I}-\boldsymbol{G^1})\boldsymbol{UR}, \mathbb{E}[(\boldsymbol{I}-\boldsymbol{G^1})\boldsymbol{G^1}\boldsymbol{y^1}(\hat{\boldsymbol{\theta}}^1_{\boldsymbol{2SLS}})|\boldsymbol{X^1}], (\boldsymbol{I}-\boldsymbol{G^1})(\boldsymbol{G^2})^2\boldsymbol{X^2}]$.

I modify 10, taking expectations given $\boldsymbol{X^2}$ and recalling $\mathbb{E}[\Delta\epsilon] = 0$:

$$(\boldsymbol{I}-\beta_2\boldsymbol{G^2})\mathbb{E}[(\boldsymbol{I}-\boldsymbol{G^1})\boldsymbol{y^2}|\boldsymbol{X^2}] = (\boldsymbol{I}-\beta_1\boldsymbol{G^1})(\boldsymbol{I}-\boldsymbol{G^1})\boldsymbol{y^1} + \tilde{\boldsymbol{\delta}}(\boldsymbol{I}-\boldsymbol{G^1})\boldsymbol{UR} +$$

$$+\gamma_2(\boldsymbol{I}-\boldsymbol{G^1})\boldsymbol{X^2_{TV}} - \gamma_1(\boldsymbol{I}-\boldsymbol{G^1})\boldsymbol{X^1_{TV}} + \delta_2(\boldsymbol{I}-\boldsymbol{G^1})\boldsymbol{G^2}\boldsymbol{X^2} - \delta_1(\boldsymbol{I}-\boldsymbol{G^1})\boldsymbol{G^1}\boldsymbol{X^1}$$

$$\mathbb{E}[(\boldsymbol{I}-\boldsymbol{G^1})\boldsymbol{y^2}|\boldsymbol{X^2}] = (\boldsymbol{I}-\beta_2\boldsymbol{G^2})^{-1}[(\boldsymbol{I}-\beta_1\boldsymbol{G^1})(\boldsymbol{I}-\boldsymbol{G^1})\boldsymbol{y^1} + \tilde{\boldsymbol{\delta}}(\boldsymbol{I}-\boldsymbol{G^1})\boldsymbol{UR} +$$

$$+\gamma_2(\boldsymbol{I}-\boldsymbol{G^1})\boldsymbol{X^2_{TV}} - \gamma_1(\boldsymbol{I}-\boldsymbol{G^1})\boldsymbol{X^1_{TV}} + \delta_2(\boldsymbol{I}-\boldsymbol{G^1})\boldsymbol{G^2}\boldsymbol{X^2} - \delta_1(\boldsymbol{I}-\boldsymbol{G^1})\boldsymbol{G^1}\boldsymbol{X^1}]$$

Let $\mathbb{E}[(\boldsymbol{I}-\boldsymbol{G^1})\boldsymbol{G^2}\boldsymbol{y^2}(\phi)|\boldsymbol{X^2},\boldsymbol{X^1}] = \boldsymbol{G^2}(\boldsymbol{I}-\beta_2\boldsymbol{G^2})^{-1}[(\boldsymbol{I}-\beta_1\boldsymbol{G^1})\mathbb{E}[(\boldsymbol{I}-\boldsymbol{G^1})\boldsymbol{y^1}(\theta^1)|\boldsymbol{X^1}] + \tilde{\boldsymbol{\delta}}(\boldsymbol{I}-\boldsymbol{G^1})\boldsymbol{UR} + \gamma_2(\boldsymbol{I}-\boldsymbol{G^1})\boldsymbol{X^2_{TV}} - \gamma_1(\boldsymbol{I}-\boldsymbol{G^1})\boldsymbol{X^1_{TV}} + \delta_2(\boldsymbol{I}-\boldsymbol{G^1})\boldsymbol{G^2}\boldsymbol{X^2} - \delta_1(\boldsymbol{I}-\boldsymbol{G^1})\boldsymbol{G^1}\boldsymbol{X^1}]$, where $\mathbb{E}[(\boldsymbol{I}-\boldsymbol{G^1})\boldsymbol{y^1}(\theta^1)|\boldsymbol{X^1}] = (\boldsymbol{I}-\beta_1\boldsymbol{G^1})^{-1}(\boldsymbol{I}-\boldsymbol{G^1})(\gamma_1\boldsymbol{I}+\delta_1\boldsymbol{G^1})\boldsymbol{X^1}$.

Then I also define the following vector $\bar{\boldsymbol{Z}} = [(\boldsymbol{I}-\boldsymbol{G^1})\boldsymbol{X^2_{TV}}, (\boldsymbol{I}-\boldsymbol{G^1})\boldsymbol{X^1_{TV}}, (\boldsymbol{I}-\boldsymbol{G^1})\boldsymbol{G^2}\boldsymbol{X^2}, (\boldsymbol{I}-\boldsymbol{G^1})\boldsymbol{G^1}\boldsymbol{X^1}, (\boldsymbol{I}-\boldsymbol{G^1})\boldsymbol{UR}$,
$\mathbb{E}[(\boldsymbol{I}-\boldsymbol{G^1})\boldsymbol{G^1}\boldsymbol{y^1}(\theta^1)|\boldsymbol{X^1}], \mathbb{E}[(\boldsymbol{I}-\boldsymbol{G^1})\boldsymbol{G^2}\boldsymbol{y^2}(\phi)|\boldsymbol{X^2},\boldsymbol{X^1}]$

I propose the following estimation procedure:

First, compute the 2SLS estimator for $\phi = (\alpha_1, \alpha_2, \beta_1, \beta_2, \gamma_1, \gamma_2, \delta_1, \delta_2)$ of the (7), using vector of instruments $\boldsymbol{M}$ and vector of covariates $\bar{\boldsymbol{X}}^{\boldsymbol{1}}$, as defined above.

$\hat{\phi}^1_{2SLS} = (\bar{\boldsymbol{X}}^T \boldsymbol{P_M} \bar{\boldsymbol{X}})^{-1} \bar{\boldsymbol{X}}^T \boldsymbol{P_M}(\boldsymbol{y^2} - \boldsymbol{y^1})$, where $\boldsymbol{P_M} = \boldsymbol{M}(\boldsymbol{M}^T \boldsymbol{M})^{-1} \boldsymbol{M}^T$ is a projection matrix.

Second, define $\hat{\bar{\boldsymbol{Z}}} = \bar{\boldsymbol{Z}}(\hat{\phi}_{2SLS}) = [(\boldsymbol{I} - \boldsymbol{G^1})\boldsymbol{X^2_{TV}}, (\boldsymbol{I} - \boldsymbol{G^1})\boldsymbol{X^1_{TV}}, (\boldsymbol{I} - \boldsymbol{G^1})\boldsymbol{G^2}\boldsymbol{X^2}, (\boldsymbol{I} - \boldsymbol{G^1})\boldsymbol{G^1}\boldsymbol{X^1}, (\boldsymbol{I} - \boldsymbol{G^1})\boldsymbol{UR}, \mathbb{E}[(\boldsymbol{I} - \boldsymbol{G^1})\boldsymbol{G^1}\boldsymbol{y^1}(\hat{\theta}^1_{2SLS})|\boldsymbol{X^1}]], \mathbb{E}[(\boldsymbol{I} - \boldsymbol{G^1})\boldsymbol{G^2}\boldsymbol{y^2}(\hat{\phi}_{2SLS})|\boldsymbol{X^2}, \boldsymbol{X^1}],$ where $\mathbb{E}[(\boldsymbol{I} - \boldsymbol{G^1})\boldsymbol{G^1}\boldsymbol{y^1}(\hat{\theta}^1_{2SLS})|\boldsymbol{X^1}] = (\boldsymbol{I} - \hat{\beta}_{1,2SLS}\boldsymbol{G^1})^{-1}(\boldsymbol{I} - \boldsymbol{G^1})(\hat{\gamma}_{1,2SLS}\boldsymbol{I} + \hat{\delta}_{1,2SLS}\boldsymbol{G^1})\boldsymbol{X^1}$, with $\hat{\theta}^1_{2SLS}$ obtained as the estimation of the first stage on the first step.

and $\mathbb{E}[(\boldsymbol{I} - \boldsymbol{G^1})\boldsymbol{G^2}\boldsymbol{y^2}(\hat{\phi}_{2SLS})|\boldsymbol{X^2}, \boldsymbol{X^1}] = \boldsymbol{G^2}(\boldsymbol{I} - \hat{\beta}_{2,2SLS}\boldsymbol{G^2})^{-1}[(\boldsymbol{I} - \hat{\beta}_{1,2SLS}\boldsymbol{G^1})\mathbb{E}[(\boldsymbol{I} - \boldsymbol{G^1})\boldsymbol{y^1}(\hat{\theta}^1_{2SLS})|\boldsymbol{X^1}] + \hat{\hat{\boldsymbol{\delta}}}_{2SLS}(\boldsymbol{I} - \boldsymbol{G^1})\boldsymbol{UR} + \hat{\gamma}_{2,2SLS}(\boldsymbol{I} - \boldsymbol{G^1})\boldsymbol{X^2_{TV}} - \hat{\gamma}_{1,2SLS}(\boldsymbol{I} - \boldsymbol{G^1})\boldsymbol{X^1_{TV}} + \hat{\delta}_{2,2SLS}(\boldsymbol{I} - \boldsymbol{G^1})\boldsymbol{G^2}\boldsymbol{X^2} - \hat{\delta}_{1,2SLS}(\boldsymbol{I} - \boldsymbol{G^1})\boldsymbol{G^1}\boldsymbol{X^1}]$

Finally, we use $\hat{\bar{\boldsymbol{Z}}}$ as a new vector of instrument to estimate (7). Then the following consistent estimator is obtained: $\hat{\phi}_{Lee} = (\hat{\bar{\boldsymbol{Z}}}^T \bar{\boldsymbol{X}})^{-1} \hat{\bar{\boldsymbol{Z}}}^T (\boldsymbol{y^2} - \boldsymbol{y^1})$.

5 Conclusion

The paper discusses the identification and estimation of the endogenous shock across the network using the newly introduced dynamic peer effect model in the presence of such shock.

I have presented the results for identification of such models, that allows disentangling the effect of unpredicted shock on the future performance. The findings of the paper suggest that it is sufficient to assume time-variability of networks together with the existence of intransitive triads (or distances of length three, depending on the correlated effects assumption) in each of the states of the network for the proposed models. Intransitive triads are present in the network if there exist two members that are only connected via the third common friend but not directly. The characteristics of friends of the friends do not influence the outcome directly, and, therefore, can be used as an instrumental variable for the friends' outcome. Such instruments can, therefore, deal with endogeneity issue. The group of new friends, different from the group of old friends, let the model capture the changes, happening due to the shock.

The procedure developed in the paper is shown to yield consistent estimators of the individual characteristics, endogenous peer effect, and the effect of unpredicted shock.

The method propose in the paper is applicable when the endogenous shocks might have the longitudinal effect on the network outcomes, such as, for example, a treatment that for some reasons cannot be randomized, or conversational networks in developing communities, etc. For example, Marchenko (2019) already tested the theoretical findings of the paper on the dataset of university students, connected via the friendship network.

References

Angelucci, Manuela and Giacomo De Giorgi (2009). "Indirect effects of an aid program: how do cash transfers affect ineligibles' consumption?" In: *The American Economic Review* 99.1, pp. 486–508.

Aral, S., L. Muchnik, and A. Sundararajan (2009). "Distinguishing influence-based contagion from homophily-driven diffusion in dynamic networks". In: *Proceedings of the National Academy of Sciences* 106.51, pp. 21544–21549.

Bramoullé, Y., H. Djebbari, and B. Fortin (2009). "Identification of peer effects through social networks". In: *Journal of econometrics* 150.1, pp. 41–55.

Calvó-Armengol, A., E. Patacchini, and Y. Zenou (2009). "Peer effects and social networks in education". In: *The Review of Economic Studies* 76.4, pp. 1239–1267.

Chandrasekhar, A. and R. Lewis (2011). "Econometrics of sampled networks". In: *Unpublished manuscript, MIT.[422].*

Comola, M. and S. Prina (2014). "Do interventions change the network? A dynamic peer effect model accounting for network changes". In: *Working paper.*

De Giorgi, Giacomo, Michele Pellizzari, and Silvia Redaelli (2010). "Identification of social interactions through partially overlapping peer groups". In: *American Economic Journal: Applied Economics*, pp. 241–275.

Dieye, Rokhaya, Habiba Djebbari, and Felipe Barrera-Osorio (2014). "Accounting for peer effects in treatment response". In:

Elhorst, J. Paul (2010). "Dynamic panels with endogenous interaction effects when T is small". In: *Regional Science and Urban Economics* 40.5, pp. 272 –282.

Goldsmith-Pinkham, P. and Guido W Imbens (2013). "Social networks and the identification of peer effects". In: *Journal of Business & Economic Statistics* 31.3, pp. 253–264.

Graham, Bryan S (2008). "Identifying social interactions through conditional variance restrictions". In: *Econometrica* 76.3, pp. 643–660.

Hinke, Stephanie von, George Leckie, and Cheti Nicoletti (2019). "The Use of Instrumental Variables in Peer Effects Models". In: *Oxford Bulletin of Economics and Statistics.*

Jackson, Matthew O. (2010). "An overview of social networks and economic applications". In: *The handbook of social economics* 1, pp. 511–585.

Kelejian, Harry H and Ingmar R Prucha (1998). "A generalized spatial two-stage least squares procedure for estimating a spatial autoregressive model with autoregressive disturbances". In: *The Journal of Real Estate Finance and Economics* 17.1, pp. 99–121.

Lee, Lung-fei (2003). "Best Spatial Two-Stage Least Squares Estimators for a Spatial Autoregressive Model with Autoregressive Disturbances". In: *Econometric Reviews* 22.4, pp. 307–335.

Li, Q., W. Zang, and L. An (2013). "Peer effects and school dropout in rural China". In: *China Economic Review* 27, pp. 238–248.

Manski, C.F. (1993). "Identification of endogenous social effects: The reflection problem". In: *The review of economic studies* 60.3, pp. 531–542.

Marchenko, M. (2019). "Endogenous Shocks in Social Networks: Exam Failures and Friends' Future Performance". In: *unpublished manuscript*.

McPherson, M., L. Smith-Lovin, and J. M. Cook (2001). "Birds of a feather: Homophily in social networks". In: *Annual review of sociology*, pp. 415–444.

Nicoletti, Cheti and Birgitta Rabe (2019). "Sibling spillover effects in school achievement". In: *Journal of Applied Econometrics* 34.4, pp. 482–501.

Paula, Áureo de (2015). "Econometrics of Network Models". In: *cemmap Working paper* CWP52/15.