Comment on "Social networks and the identification of peer effects" by Paul Goldsmith-Pinkham and Guido W. Imbens

Bryan S. Graham*

January 10, 2013

Consider a group of three, potentially connected, individuals ($i = 1, 2, 3$). Available is a large random sample of such groups. For each group sampled we observe all social ties among its constituent members in each $t = 0, 1, 2, 3$ periods. Let $D_{ijt} = 1$ if individual i is 'friends' (i.e., connected) with individual j in period t and zero otherwise. Ties are undirected so that $D_{ij} = D_{ji}$ for $i \neq j$. We rule out self-ties so that $D_{ii} = 0$. The network adjacency matrix in period t is denoted by $\mathbf{D}_t$ with typical element D_{ijt}. The sampling process asymptotically reveals

$$f(\mathbf{d}_3, \mathbf{d}_2, \mathbf{d}_1, \mathbf{d}_0) = \Pr(\mathbf{D}_3 = \mathbf{d}_3, \mathbf{D}_2 = \mathbf{d}_2, \mathbf{D}_1 = \mathbf{d}_1, \mathbf{D}_0 = \mathbf{d}_0).$$

Let $F_{ijt} = 1$ if i and j have any friends in common during period t and zero otherwise. For example if i and k, as well as j and k, are connected in period t, then i and j will share the common friend k such that $F_{ijt} = 1$. Note that i and j need not be direct friends themselves.

Let ν_i and ξ_i be individual-specific, time-invariant, latent variables. The latent social distance between individuals i and j is measured by the distance function $g(\xi_i, \xi_j)$. This distance function (i) takes a value of zero if $\xi_i = \xi_j$, (ii) is symmetric in its arguments, and (iii) is increasing in $|\xi_i - \xi_j|$. Let the pair-specific unobserved heterogeneity term $A_{ij} = \nu_i + \nu_j - g(\xi_i, \xi_j) = A_{ji}$.

Let $\mathbf{1}(\bullet)$ denote the indicator function; individuals i and j form a link in periods t=1,2,3

*Department of Economics, University of California - Berkeley and the National Bureau of Economic Research, $\mathtt{e-mail:}$ bgraham@econ.berkeley.edu, $\mathtt{web:}$ http://elsa.berkeley.edu/ bgraham/. These comments were prepared for the *Journal of Business and Economic Statistics* Invited Address delivered by Guido Imbens on January 7th at the 2012 annual meeting of the American Economic Association. I am grateful to Kei Hirano and Jonathan Wright, in their capacity as co-editors, for the opportunity to comment on Guido Imbens' lecture.

according to the rule

$$D_{ijt} = \mathbf{1}\left(\alpha + \beta D_{ijt-1} + \gamma F_{ijt-1} + A_{ij} - U_{ijt} > 0\right). \tag{1}$$

The term inside the indicator function in (1) is the net social surplus associated with a link between i and j. Agents form a link if the net utility from doing so is positive. Goldsmith-Pinkham and Imbens (2012) model utility at the individual level and require positivity of both candidate partner utilities in order for a link to form. If utility is transferable rule (1) seems reasonable (see Graham, 2011). When utility is non-transferable the approach of Goldsmith-Pinkham and Imbens is most appropriate.

The match-by-period specific utility shock, U_{ijt}, is independently and identically distributed across pairs and over time with distribution function $F(u)$. This distribution function may vary arbitrarily with $\mathbf{A} = (A_{12}, A_{13}, A_{23})'$, but I suppress any such dependence in what follows.[1]

An important feature of (1) is that it is backwards-looking. This eliminates contemporaneous feedback which can complicate identification (cf., Manski, 1993). While i and j's decision to link or not *will* be influenced by the past link history of, say, agent k, it *will not* be influenced by any link decisions involving agent k in the current period.

Model (1) incorporates four types of network dependencies emphasized in prior research (cf., Snijders, 2011). First, the $-g(\xi_i, \xi_j)$ component of A_{ij} in (1) increases the probability of ties across similar individuals. All other things equal, agents will assortatively match on ξ_j. Call this effect *homophily* (McPherson, Smith-Lovin and Cook, 2001; Jackson, 2008). Goldsmith-Pinkham and Imbens (Section 7) choose $g(\xi_i, \xi_j) = \alpha_\xi |\xi_i - \xi_j|$ for their empirical model of network formation. Second, the presence of ν_i and ν_j in A_{ij} allows for *degree heterogeneity* (cf., Krivitsky, Handcock, Raftery and Hoff, 2009). If ν_i is high, then the net surplus associated with *any* match involving i will tend to be high (e.g., i might be a 'good friend' generically). Goldsmith-Pinkham and Imbens do not incorporate degree heterogeneity into their model, but such heterogeneity appears to be an important feature of real world networks. Third the presence of F_{ijt-1} in (1) implies a taste for *transitivity* in link formation or 'triadic closure'. Specifically a link between i and j in the current period is, all things equal, more likely if they shared a common friend in the prior period. The strength of transitivity in link formation is indexed by the parameter γ. Finally, the parameter β captures *state-dependence* in ties: i and j are more likely to be friends in period t if they were friends in period $t-1$.

[1] More precisely each sampled network may have its own distribution for U_{ijt}. However the assumption that U_{ijt} is independently and identically distributed across potential ties and over time *within* a network is essential.

Discriminating between homophily/degree heterogeneity and transitivity in network formation is scientifically interesting and important from the perspective of the policy-maker. To motivate this assertion consider the following stylized example. Let groups correspond to schools and ξ_i an index of socioeconomic background. Due to the sorting of families across neighborhoods, the distribution of ξ_i within-schools is likely to be considerably more compressed than that between-schools. Consequently we may observe, for reasons of homophily, a large number of triangles (i.e., networks where agents 1, 2 and 3 are all connected) across our sample of schools. A preponderance of triangles may occur even in the absence of any structural taste for transitivity in links. The structural source of clustering is nevertheless policy-relevant. In the presence of transitivity a teacher or principle may be able to alter the network structure within a school by facilitating tie-formation or -dissolution across a small number of students.[2] In the absence of a taste for transitivity such manipulations may be much more difficult to engineer.

While the *homophily* versus *transitivity* identification problem has been informally articulated in the literature on network formation (e.g., Goodreau, Kitts and Morris, 2009; Kitts and Huang, 2011; Miyauchi, 2012), I am aware of no systematic analysis of it.

The joint density attached to a specific realization of a sequence of network structures and social distances is given by

$$\begin{aligned} f\left(\mathbf{d_3},\mathbf{d}_2,\mathbf{d}_1,\mathbf{d}_0,\mathbf{a}\right) &= \prod_{t=1}^{3}\Pr\left(\mathbf{D}_t=\mathbf{d}_t\middle|\,\mathbf{D_{t-1}}=\mathbf{d_{t-1}},\mathbf{A}=\mathbf{a}\right) \\ &\quad \times\pi\left(\mathbf{d_0},\mathbf{a}\right) \\ &= \prod_{t=1}^{3}\prod_{\{i<j,i\neq k,j\neq k\}}\Big\{F\left(\alpha+\beta d_{ijt-1}+\gamma d_{ikt-1}d_{jkt-1}+a_{ij}\right)^{d_{ijt}} \\ &\quad \times\left[1-F\left(\alpha+\beta d_{ijt-1}+\gamma d_{ikt-1}d_{jkt-1}+a_{ij}\right)\right]^{1-d_{ijt}}\Big\} \\ &\quad \times\pi\left(\mathbf{d_0},\mathbf{a}\right) \end{aligned} \tag{2}$$

where $\pi\left(\mathbf{d_0},\mathbf{a}\right)$ is the joint density of the initial network structure and vector of pair-specific heterogeneity terms (I have used the fact that $F_{ijt}=D_{ikt}D_{jkt}$). Goldsmith-Pinkham and Imbens assume independence of $\mathbf{D_0}$ and $\mathbf{A}$. If the network under consideration began prior to the initial period of observation this assumption seems implausible. The link rule given in (1) induces dependence between network structure and $\mathbf{A}$ in later periods. The issues involved are related to those of the initial conditions problem in dynamic binary choice panel data models (e.g., Heckman, 1981a-c). Goldsmith-Pinkham and Imbens further model the ξ_i as independent draws from a two mass-point distribution with known mixing probabilities. This

[2] In the education context the manipulation of social cliques may be of special interest.

assumption in turn induces a distribution for $\mathbf{A}$. While their approach does result in some dependence across the elements of $\mathbf{A}$, it is of a highly structured form. The stylized example given above suggests that the dependence structure across the elements of $\mathbf{A}$ may be quite complex.[3] Experience drawn from the literature on discriminating between state-dependence and heterogeneity using panel data suggests that Goldsmith-Pinkham and Imbens' modeling assumptions lead them to overstate the role of past network structure in explaining current link formation. This is, of course, only a conjecture. It nevertheless motivates the question of what can be learned without imposing strong assumptions on the form of $\pi(\mathbf{d_0}, \mathbf{a})$?

Here I wish to formally study identification of β and, especially, γ. In contrast to Goldsmith-Pinkham and Imbens my treatment will be 'fixed effects' in nature – the joint distribution of $\mathbf{D}_0$ and $\mathbf{A}$ will be left unrestricted. Link decisions across pairs of agents are conditionally independent given the network structure in the prior period and the *unobserved* vector of latent social distances $\mathbf{A}$. However, unconditional on $\mathbf{A}$ the dependence across different link decisions is allowed to be quite complex. There is some connection between my question and that of identifying state-dependence using panel data (e.g., Chamberlain, 1985). However the analogy is incomplete; a network structure is induced by the inter-connected decisions of multiple agents. Modeling a sequence of networks is considerably harder than modeling a sequence of discrete decisions made by a single agent.

As a preliminary analysis I will study what can be learned by the observed sequence of link decisions between agents i and j conditional on the observed sequence of links between the remaining agents. The goal is to find features of (2) that do not dependent on $\mathbf{A}$.

Without loss of generality set $i = 1$ and $j = 3$ and consider what can be learned from the relative frequency of observing a specific member of the set of network histories

$$\begin{aligned}\mathbf{E}_{01} &= \left\{ \mathbf{D}_0 = \begin{pmatrix} 0 & D_{120} & D_{130} \\ D_{120} & 0 & D_{230} \\ D_{130} & D_{230} & 0 \end{pmatrix}, \mathbf{D}_1 = \begin{pmatrix} 0 & D_{121} & 0 \\ D_{121} & 0 & D_{231} \\ 0 & D_{231} & 0 \end{pmatrix}, \right. \\ & \left. \mathbf{D}_2 = \begin{pmatrix} 0 & D_{121} & 1 \\ D_{121} & 0 & D_{231} \\ 1 & D_{231} & 0 \end{pmatrix}, \mathbf{D}_3 = \begin{pmatrix} 0 & D_{121} & D_{133} \\ D_{121} & 0 & D_{231} \\ D_{133} & D_{231} & 0 \end{pmatrix} \right\}\end{aligned}$$

[3]Specifically, cross group sorting suggests that ξ_1, ξ_2 and ξ_3 covary, which, in turn, induces dependence across the elements of $\mathbf{A}$.

versus a member of the set

$$\mathbf{E}_{10} = \left\{ \mathbf{D}_0 = \begin{pmatrix} 0 & D_{120} & D_{130} \\ D_{120} & 0 & D_{230} \\ D_{130} & D_{230} & 0 \end{pmatrix}, \mathbf{D}_1 = \begin{pmatrix} 0 & D_{121} & 1 \\ D_{121} & 0 & D_{231} \\ 1 & D_{231} & 0 \end{pmatrix}, \right.$$

$$\left. \mathbf{D}_2 = \begin{pmatrix} 0 & D_{121} & 0 \\ D_{121} & 0 & D_{231} \\ 0 & D_{231} & 0 \end{pmatrix}, \mathbf{D}_3 = \begin{pmatrix} 0 & D_{121} & D_{133} \\ D_{121} & 0 & D_{231} \\ D_{133} & D_{231} & 0 \end{pmatrix} \right\}.$$

The networks contained in the sets $\mathbf{E}_{01}$ and $\mathbf{E}_{10}$ have two key features. First, any $(i,k) = (1,2)$ and $(k,j) = (2,3)$ ties are stable across periods 1, 2 and 3. If either agent i or j is linked to k in period 1, they are also linked in periods 2 and 3. Likewise the absence of a period 1 link is associated with an absence of a period 2 and 3 link. The (i,k) and (k,j) pairs may switch their link status between periods 0 and 1, and it is essential that this occurs in some sampled networks, but they do not revise their link status in subsequent periods. This last feature of $\mathbf{E}_{01}$ and $\mathbf{E}_{10}$ ensures variation over time in the opportunity for agents i and j to engineer triadic closure by forming a link. We can say that the (i,j) is embedded in a stable neighborhood. Second, the sequence of (i,j) links differs across $\mathbf{E}_{01}$ and $\mathbf{E}_{10}$. In the first set of histories (i,j) are linked in period 2, but not 1, while in the second this ordering is reversed. We require that agents i and j switch their link status between periods 1 and 2.

The sets $\mathbf{E}_{01}$ and $\mathbf{E}_{10}$ were selected by a combination of educated guessing, inspired by the work of Cox (1954), Chamberlain (1985), and Honoré and Kyriazidou (2000) on the identification of dynamic binary choice panel data models, and trial and error. Let $\mathbf{D} = (\mathbf{D}_3, \mathbf{D}_2, \mathbf{D}_1, \mathbf{D}_0)$ denote the full sequence of network structures and $\mathbf{e}_{01}$ a specific element of the set $\mathbf{E}_{01}$ (and similarly for $\mathbf{e}_{10}$). A straightforward, albeit tedious, calculation gives (see the Appendix)

$$\frac{\Pr\left(\mathbf{D} = \mathbf{e}_{01} \middle| \mathbf{A} = \mathbf{a}\right)}{\Pr\left(\mathbf{D} = \mathbf{e}_{10} \middle| \mathbf{A} = \mathbf{a}\right)} = \frac{1 - F\left(\alpha + \beta d_{130} + \gamma d_{120} d_{230} + a_{13}\right)}{F\left(\alpha + \beta d_{130} + \gamma d_{120} d_{230} + a_{13}\right)} \frac{F\left(\alpha + \beta d_{133} + \gamma d_{121} d_{231} + a_{13}\right)}{1 - F\left(\alpha + \beta d_{133} + \gamma d_{121} d_{231} + a_{13}\right)}.$$

Monotonicity of F then implies that (cf., Manski, 1987; Honoré and Kyriazidou, 2000; Graham, 2011, Section 4.3):

$$\operatorname{sgn}\left(\Pr\left(\mathbf{E}_{01} = \mathbf{e}_{01} \middle| \mathbf{a}\right) - \Pr\left(\mathbf{E}_{10} = \mathbf{e}_{10} \middle| \mathbf{a}\right)\right) = \operatorname{sgn}\left(\beta\left(d_{133} - d_{130}\right) + \gamma\left(d_{121} d_{231} - d_{120} d_{230}\right)\right). \tag{3}$$

By separately considering the subset of networks with, respectively $d_{133} \neq d_{130}$ and $d_{121}d_{231} = d_{120}d_{230}$ and $d_{133} = d_{130}$ and $d_{121}d_{231} \neq d_{120}d_{230}$, we can show that the signs of β and γ are separately identified. Consequently the presence of transitivity is identifiable without imposing any restrictions on the joint distribution of $\mathbf{D}_0$ and $\mathbf{A}$. To my knowledge this is a new result. In ongoing work I have shown that this result may be extended to networks of arbitrary size (Graham, 2012).

If we additionally assume that U_{ijt} is logistically distributed, as in Goldsmith-Pinkham and Imbens (2011), we have identification up to scale with

$$\Pr\left(\mathbf{D} = \mathbf{e_{01}} \middle| \mathbf{A} = \mathbf{a}, \mathbf{D} \in \{\mathbf{e}_{01}, \mathbf{e}_{10}\}\right) = \frac{\exp\left(\frac{\beta(d_{133}-d_{130})+\gamma(d_{121}d_{231}-d_{120}d_{230})}{\sigma}\right)}{1+\exp\left(\frac{\beta(d_{133}-d_{130})+\gamma(d_{121}d_{231}-d_{120}d_{230})}{\sigma}\right)}, \qquad (4)$$

where σ is the scale parameter for U_{ijt} (typically normalized to 1).

That, in the context of a simple dynamic model of network formation, it is possible to discriminate between transitivity and homophily/degree heterogeneity while leaving the joint distribution of $\mathbf{D}_0$ and $\mathbf{A}$ unrestricted is an encouraging result. Goldsmith-Pinkham and Imbens make very strong assumptions on this joint distribution and I suspect their conclusions may be sensitive to them.

While I have presented a simple 'fixed effect' identification result, I remain sympathetic to the '(correlated) random effects' modeling strategy of Goldsmith-Pinkham and Imbens (i.e., an approach that does impose restrictions on the joint distribution of $\mathbf{D}_0$ and $\mathbf{A}$). As in nonlinear panel data analysis the two approaches are complementary. Goldsmith-Pinkham and Imben's application, however, indicates that working with the integrated likelihood is numerically challenging. Some of their choices regarding the joint distribution of $\mathbf{D}_0$ and $\mathbf{A}$ appear to be driven by computational concerns. It would be useful to construct parsimonious parametric models for this distribution that incorporate richer forms of dependence. Of course, continued study of 'fixed effects' approaches is also warranted.

A Appendix

We begin by evaluating (2) at $\mathbf{E}_{01} = \mathbf{e}_{01}$ and $\mathbf{E}_{10} = \mathbf{e}_{10}$. We get

$$
\begin{aligned}
\Pr\left(\mathbf{D} = \mathbf{e}_{01} \middle| \mathbf{A} = \mathbf{a}\right) \;=\; & \pi\left(\mathbf{d_0}, \mathbf{a}\right) \\
& \times F\left(\alpha + \beta d_{120} + \gamma d_{130} d_{230} + a_{12}\right)^{d_{121}} \left[1 - F\left(\alpha + \beta d_{120} + \gamma d_{130} d_{230} + a_{12}\right)\right]^{1-d_{121}} \\
& \times \left[1 - F\left(\alpha + \beta d_{130} + \gamma d_{120} d_{230} + a_{13}\right)\right] \\
& \times F\left(\alpha + \beta d_{230} + \gamma d_{120} d_{130} + a_{23}\right)^{d_{231}} \left[1 - F\left(\alpha + \beta d_{230} + \gamma d_{120} d_{130} + a_{23}\right)\right]^{1-d_{231}} \\
& \times F\left(\alpha + \beta d_{121} + a_{12}\right)^{d_{121}} \left[1 - F\left(\alpha + \beta d_{121} + a_{12}\right)\right]^{1-d_{121}} \\
& \times F\left(\alpha + \gamma d_{121} d_{231} + a_{13}\right) \\
& \times F\left(\alpha + \beta d_{231} + a_{23}\right)^{d_{231}} \left[1 - F\left(\alpha + \beta d_{231} + a_{23}\right)\right]^{1-d_{231}} \\
& \times F\left(\alpha + \beta d_{121} + \gamma d_{231} + a_{12}\right)^{d_{121}} \left[1 - F\left(\alpha + \beta d_{121} + \gamma d_{231} + a_{12}\right)\right]^{1-d_{121}} \\
& \times F\left(\alpha + \beta + \gamma d_{121} d_{231} + a_{13}\right)^{d_{133}} \left[1 - F\left(\alpha + \beta + \gamma d_{121} d_{231} + a_{13}\right)\right]^{1-d_{133}} \\
& \times F\left(\alpha + \beta d_{231} + \gamma d_{121} + a_{23}\right)^{d_{231}} \left[1 - F\left(\alpha + \beta d_{231} + \gamma d_{121} + a_{23}\right)\right]^{1-d_{231}}
\end{aligned}
$$

and

$$
\begin{aligned}
\Pr\left(\mathbf{D} = \mathbf{e}_{10} \middle| \mathbf{A} = \mathbf{a}\right) \;=\; & \pi\left(\mathbf{d_0}, \mathbf{a}\right) \\
& \times F\left(\alpha + \beta d_{120} + \gamma d_{130} d_{230} + a_{12}\right)^{d_{121}} \left[1 - F\left(\alpha + \beta d_{120} + \gamma d_{130} d_{230} + a_{12}\right)\right]^{1-d_{121}} \\
& \times F\left(\alpha + \beta d_{130} + \gamma d_{120} d_{230} + a_{13}\right) \\
& \times F\left(\alpha + \beta d_{230} + \gamma d_{120} d_{130} + a_{23}\right)^{d_{231}} \left[1 - F\left(\alpha + \beta d_{230} + \gamma d_{120} d_{130} + a_{23}\right)\right]^{1-d_{231}} \\
& \times F\left(\alpha + \beta d_{121} + \gamma d_{231} + a_{12}\right)^{d_{121}} \left[1 - F\left(\alpha + \beta d_{121} + \gamma d_{231} + a_{12}\right)\right]^{1-d_{121}} \\
& \times \left[1 - F\left(\alpha + \beta + \gamma d_{121} d_{231} + a_{13}\right)\right] \\
& \times F\left(\alpha + \beta d_{231} + \gamma d_{121} + a_{23}\right)^{d_{231}} \left[1 - F\left(\alpha + \beta d_{231} + \gamma d_{121} + a_{23}\right)\right]^{1-d_{231}} \\
& \times F\left(\alpha + \beta d_{121} + a_{12}\right)^{d_{121}} \left[1 - F\left(\alpha + \beta d_{121} + a_{12}\right)\right]^{1-d_{121}} \\
& \times F\left(\alpha + \gamma d_{121} d_{231} + a_{13}\right)^{d_{133}} \left[1 - F\left(\alpha + \gamma d_{121} d_{231} + a_{13}\right)\right]^{1-d_{133}} \\
& \times F\left(\alpha + \beta d_{231} + a_{23}\right)^{d_{231}} \left[1 - F\left(\alpha + \beta d_{231} + a_{23}\right)\right]^{1-d_{231}}
\end{aligned}
$$

Taking the ratio of these two probabilities yields, after some obvious cancellations:

$$\begin{aligned}\frac{\Pr\left(\mathbf{D}=\mathbf{e}_{01}\right|\mathbf{A}=\mathbf{a})}{\Pr\left(\mathbf{D}=\mathbf{e}_{10}\right|\mathbf{A}=\mathbf{a})} &= \frac{1-F\left(\alpha+\beta d_{130}+\gamma d_{120}d_{230}+a_{13}\right)}{F\left(\alpha+\beta d_{130}+\gamma d_{120}d_{230}+a_{13}\right)} \\ &\times\frac{F\left(\alpha+\gamma d_{121}d_{231}+a_{13}\right)}{1-F\left(\alpha+\beta+\gamma d_{121}d_{231}+a_{13}\right)} \\ &\times\frac{F\left(\alpha+\beta+\gamma d_{121}d_{231}+a_{13}\right)^{d_{133}}\left[1-F\left(\alpha+\beta+\gamma d_{121}d_{231}+a_{13}\right)\right]^{1-d_{133}}}{F\left(\alpha+\gamma d_{121}d_{231}+a_{13}\right)^{d_{133}}\left[1-F\left(\alpha+\gamma d_{121}d_{231}+a_{13}\right)\right]^{1-d_{133}}}.\end{aligned}$$

Now observe that if $d_{133}=1$ we have $F\left(\alpha+\beta+\gamma d_{121}d_{231}+a_{13}\right)^{d_{133}}=F\left(\alpha+\beta d_{133}+\gamma d_{121}d_{231}+a_{13}\right)$ and $1-F\left(\alpha+\beta+\gamma d_{121}d_{231}+a_{13}\right)=1-F\left(\alpha+\beta d_{133}+\gamma d_{121}d_{231}+a_{13}\right)$ which implies the simplification given immediately prior to equation (3) of the main text. Similarly if $d_{133}=0$ we have $\left[1-F\left(\alpha+\gamma d_{121}d_{231}+a_{13}\right)\right]^{1-d_{133}}=\left[1-F\left(\alpha+\beta d_{133}+\gamma d_{121}d_{231}+a_{13}\right)\right]$ and $F\left(\alpha+\gamma d_{121}d_{231}+a_{13}\right)=F\left(\alpha+\beta d_{133}+\gamma d_{121}d_{231}+a_{13}\right)$ which gives the same result.

In the logistic case we have

$$\frac{\frac{1-F(\alpha+\beta d_{130}+\gamma d_{120}d_{230}+a_{13})}{F(\alpha+\beta d_{130}+\gamma d_{120}d_{230}+a_{13})}}{\frac{1-F(\alpha+\beta d_{133}+\gamma d_{121}d_{231}+a_{13})}{F(\alpha+\beta d_{133}+\gamma d_{121}d_{231}+a_{13})}} = \exp\left(\frac{\beta\left(d_{133}-d_{130}\right)+\gamma\left(d_{121}d_{231}-d_{120}d_{230}\right)}{\sigma}\right),$$

from which (4) follows directly.

References

[1] Chamberlain, Gary. (1985). "Heterogeneity, omitted variable bias, and duration dependence," *Longitudinal Analysis of Labor Market Data*: 3 - 38 (J.J. Heckman & B. Singer, Eds.). Cambridge: Cambridge University Press.

[2] Chamberlain, Gary. (1993). "Feedback in panel data models," *Mimeo.*

[3] Cox, D. R. (1958). "The regression analysis of binary sequences," *Journal of the Royal Statistical Society B* 20 (2): 215 - 241.

[4] Goodreau, Steven M., James A. Kitts and Martina Morris. (2009). "Birds of a feather, or friend of a friend? Using exponential random graph models to investigate adolescent social networks," *Demography* 46 (1): 103 - 125.

[5] Graham, Bryan S. (2011). "Econometric methods for the analysis of assignment problems in the presence of complementarity and social spillovers," *Handbook of Social Economics* 1B: 965 - 1052 (J. Benhabib, A. Bisin, & M. Jackson, Eds.). Amsterdam: North-Holland.

[6] Graham, Bryan S. (2012). "Homophily and transitivity in dynamic network formation," *Mimeo.*

[7] Hahn, Jinyong. (2001). "The information bound of a dynamic panel logit model with fixed effects," *Econometric Theory* 17 (5): 913 - 932.

[8] Heckman, James. J. (1981a) "Heterogeneity and state dependence," *Studies in Labor Markets*: 91 - 139 (S. Rosen, Ed.). Chicago: University of Chicago Press.

[9] Heckman, James. J. (1981b). "Statistical models for discrete panel data," *Structural Analysis of Discrete Data and Econometric Applications*: 114 - 178 (C.F. Manski & D.L. McFadden, Eds.). Cambridge, MA: The MIT Press.

[10] Heckman, James. J. (1981c). "The incidental parameters problem and the problem of initial conditions in estimating a discrete time-discrete data stochastic process," *Structural Analysis of Discrete Data and Econometric Applications*: 179 - 195 (C.F. Manski & D.L. McFadden, Eds.). Cambridge, MA: The MIT Press.

[11] Honoré, Bo E. and Ekaterini Kyriazidou. (2000). "Panel data discrete choice models with lagged dependent variables," *Econometrica* 68 (4): 839 - 874.

[12] Imbens, Guido W. and Paul Goldsmith-Pinkham. (2011). "Social networks and the identification of peer effects," *Mimeo.*

[13] Jackson, Matthew O. (2008). *Social and Economic Networks.* Princeton, NJ: Princeton University Press.

[14] Kitts, James A. and Jun Huang. (2011). "Triads," *Encyclopedia of Social Networks* 2: 873 - 874 (G. Barnett, Ed.). New York: Sage Publications, Inc.

[15] Krivitsky, Pavel N., Mark S. Handcock, Adrian E. Raftery, and Peter D. Hoff. (2009). "Representing degree distributions, clustering, and homophily in social networks with latent cluster random effects models," *Social Networks* 31 (3): 204 - 213.

[16] Manski, Charles F. (1987). "Semiparametric analysis of random effects linear models from binary panel data," *Econometrica* 55 (2): 357 - 362.

[17] Manski, Charles F. (1993). "Identification of endogenous social effects: the reflection problem," *Review of Economic Studies* 60 (3): 531 - 542.

[18] McPherson, Miller, Lynn Smith-Lovin and James M. Cook. (2001). "Birds of a feather: homophily in social networks," *Annual Review of Sociology* 27 (1): 415 - 444.

[19] Miyauchi, Yuhei. (2012). "Structural estimation of a pairwise stable network formation with nonnegative externality," *Mimeo.*

[20] Snijders, Tom A.B. (2011). "Statistical models for social networks," *Annual Review of Sociology* 37 (1): 131 - 153.