Homophily in preferences or meetings? Identifying and estimating an iterative network formation model*

Luis Antonio Fantozzi Alvarez† Cristine Campos de Xavier Pinto‡
Vladimir Pinheiro Ponczek§

January 20, 2020

Abstract

Is homophily in social and economic networks driven by a taste for homogeneity (preferences) or by a higher probability of meeting individuals with similar attributes (opportunity)? This paper studies identification and estimation of an iterative network game that distinguishes between both mechanisms. We provide conditions under which preference- and meeting-related parameters are identified. We then analyse estimation under both frequentist and Bayesian perspectives. As an application, we study the role of preferences and meetings in shaping intraclassroom friendship networks in Northeastern Brazil. We find that tracking students may lead to welfare improvements, though the benefit diminishes over the school year.

Keywords: homophily; network formation; identification at infinity; lilelihood-free estimation.

*We would like to thank Angelo Mele, Áureo de Paula, Braz Camargo, Bruno Ferman, Eduardo Mendes, Ricardo Masini and Sérgio Firpo for their useful comments and sugestions. We are also grateful to seminar participants at PUC-Rio and the 41st SBE meeting.

†Doctoral student in Statistics, IME/USP. E-mail: alvarez@ime.usp.br.

‡Associate Professor of Economics, EESP/FGV.

§Associate Professor of Economics, EESP/FGV.

1 Introduction

Homophily, the observed tendency of agents with similar attributes maintaining relationships, is a salient feature in social and economic networks (Chandrasekhar, 2016; Jackson, 2010, and references therein). Inasmuch as it may drive network formation, homophily can produce nonnegligible effects in outcomes as diverse as smoking behaviour (Badev, 2017) and test scores (Hsieh and Lee, 2016; Goldsmith-Pinkham and Imbens, 2013). Perhaps unsurprisingly then, the appropriate modelling of homophily has received quite a deal of attention in the recent push for estimable econometric models of network formation (Goldsmith-Pinkham and Imbens, 2013; Chandrasekhar and Jackson, 2016; Mele, 2017; Graham, 2016, 2017).

A strand of the literature on the topic makes a distinction between homophily that is due to choice; and homophily that is due to opportunity[1]. We shall label the former homophily "in preferences"; and the latter homophily in "meetings". This distinction has not only theoretical, but also practical appeal: whereas public policy may be able to alter the meeting technology between agents (say, by desegregating environments), it may be less successful in changing preferences. Theoretical models that distinguish between these mechanisms can be found in Currarini et al. (2009) and Bramoullé et al. (2012)[2]. These models have some limitations, though: first, they focus on steady-state or "long-run" behaviour, which may not be appropriate in settings where transitional dynamics may matter (as in our applied section); second, they are either purely probabilistic[3] (Bramoullé et al., 2012) or, in the case of Currarini et al. (2009), allow for only a restrictive set of pay-offs from relationships[4]. These limitations render these models unfit for some empirical analyses.

This study aims to fill in the gap by providing an estimable econometric model that accounts for both homophily in "preferences" and "meetings". We study identification and

[1] Cf. Jackson (2010, p. 68) for a discussion.

[2] See also Currarini et al. (2010) for an attempt at estimation.

[3] In contrast to strategic models of network formation. See Jackson (2010) and de Paula (2016) for examples.

[4] Pay-offs in Currarini et al. (2009) depend only on the number of relationships with individuals of the same or different types. There is no role for indirect benefits.

estimation of a sequential network formation algorithm originally found in Mele (2017) (also Christakis et al. (2010) and Badev (2017)), where agents meet sequentially in pairs in order to revise their relationship status. The model is well-grounded in the theoretical literature of strategic network formation (Jackson and Watts, 2002) and allows specifications that account for both "homophilies". In spite of that, our approach differs from previous work in the literature in several aspects. While Mele (2017) discusses identification and estimation of utility parameters off from the model's induced stationary distribution under large- and many-network asymptotics, we study identification and estimation of *both* preference- and meeting-related parameters under many-network asymptotics in a setting where networks are observed in two points of time[5]. Our results also cover a larger – and arguably less restrictive – class of pay-offs and meeting processes than those in Mele (2017), where the author assumes utilities admit a potential function and meeting probabilities do not depend on the existence of a link between agents in the current network. By studying identification and estimation of general classes of preference- and matching-related parameters in (possibly) off-stationary-equilibrium settings, we hope to contribute to the model's applicability and empirical usefulness, especially in conducting counterfactual analyses.

As an application, we study how "homophilies" structure network formation in primary schools in Northeastern Brazil (Pinto and Ponczek, 2017). We consider 30 municipal elementary schools in Recife, Pernambuco, for which baseline (early 2014) and followup (late 2014) data on 3rd- and 5th- grade intra-classroom friendship networks was collected. Using this information, we are able to assess how changes in the meeting technology between classmates impact homophily in friendships. Our results suggest that removing biases in meeting opportunities would actually increase observed homophily patterns in students' cognitive skills. We also use our estimated model in welfare assessments. We find that unbiased matching leads to a lower aggregate utility path across the school year; and that tracking students according to their cognitive skills leads to welfare improvements, though this benefit appears to diminish in the long run.

In the next sections, we introduce the network formation game under consideration (Section 2); explore identification when information on the network structure is available

[5]In fact, as we show in Appendix B, meeting parameters are unidentified in the setting of Mele (2017).

in two distinct points of time (Section 3); and discuss estimation under both a frequentist and a Bayesian perspective (Section 4). Section 5 presents the results of our application. Section 6 concludes.

2 Setup

The setup expands upon Mele (2017). We consider a network game on a finite set of agents $\mathcal{I} \coloneqq \{1, 2 \ldots N\}$. Each agent $i \in \mathcal{I}$ is endowed with a $k \times 1$ vector of exogenous characteristics W_i. Agents' vectors are stacked on matrix $X \coloneqq \begin{bmatrix} W_1 & W_2 & \cdots & W_N \end{bmatrix}'$. Agents' characteristics are drawn according to law $\mathbb{P}_X$ before the game starts and remain *fixed* throughout. We denote the support of X by $\mathcal{X}$ and a realisation of X by an element $x \in \mathcal{X}$.

Time is discrete. At each round $t \in \mathbb{N}$ of the network formation process, agents' relations are described by a *directed network*. Information on the network is stored on a $N \times N$ *adjacency matrix*, with entry $g_{ij} = 1$ if i lists j as a friend and 0 otherwise[6]. By assumption, $g_{ii} = 0$ for all $i \in \mathcal{I}$. We denote the set of all $2^{N(N-1)}$ possible adjacency matrices by $\mathcal{G}$.

Agent i's utility from a network g when covariates are $X = x$ is described by a utility function $u_i : \mathcal{G} \times \mathcal{X} \mapsto \mathbb{R}$, $u_i(g, x)$. Utility may depend on the entire network and on the entire set of agents' covariates.

Agents are myopic, i.e. they form, maintain or sever relationships based on the current utility these bring. At each round, a *matching process* m^t selects a pair of agents (i, j). A matching process m^t is a stochastic process $\{m^t : t \in \mathbb{N}\}$ over $\mathcal{M} \coloneqq \{(i, j) \in \mathcal{I} \times \mathcal{I} : i \neq j\}$. If the pair (i, j) is selected, agent i will get to choose whether to form/mantain or not form/sever a relationship with j. After the matching process selects a pair of agents, a pair of choice-specific idiosyncratic shocks $(\epsilon_{ij,t}(0), \epsilon_{ij,t}(1))$ are drawn, where $\epsilon_{ij,t}(1)$ corresponds to the taste shock in forming/maintaining a relationship with j at time t. These shocks are unobserved by the econometrician and enter additively in the utility of each choice[7].

[6]Our model and main results are easily extended to the case of an *undirected* network where friendships are forcibly symmetric.

[7]Additive separability of unobserved shocks is a common assumption in the econometric literature on

Given that choice is myopic, agent i forms/maintain a relation with j iff:

$$u_i([1, g_{-ij}], X) + \epsilon_{ij,t}(1) \geq u_i([0, g_{-ij}], X) + \epsilon_{ij,t}(0) \tag{1}$$

where $[a, g_{-ij}]$ denotes an adjacency matrix with all entries equal to g except for entry ij, which equals a.

The next two assumptions constrain the meeting process and the distribution of shocks.

Assumption 2.1. The meeting process $\{m^t : t \in \mathbb{N}\}$ is described by a *time-invariant* matching function $\rho : \mathcal{M} \times \mathcal{G} \times \mathcal{X} \mapsto [0, 1]$, where $\rho((i, j), g, x)$ is the probability that (i, j) is selected when covariates are $X = x$ and *the previous-round network was* g. Moreover, for all $g \in \mathcal{G}$, $x \in \mathcal{X}$, $(i, j) \in \mathcal{M}$, $\rho((i, j), g, x) > 0$.

So the matching function assigns positive probability to all possible meetings under all possible values of covariates and previous-round networks. Note that this allows for dependence on the existence of previous-round links, which was not permitted in Mele (2017)[8].

Assumption 2.2. Shocks are drawn iid across pairs and time, indenpendently from X, from a *known* distribution $(\epsilon(0), \epsilon(1))' \sim F_\epsilon$ which is absolutely continuous with respect to the Lebesgue measure on $\mathbb{R}^2$ and has positive density almost everywhere.

Conditional on $X = x$, we have that, under Assumption 2.1 and Assumption 2.2 – and specifying some initial distribution $\mu_0(x) \in \Delta(\mathcal{G})$[9] –, the network game just described induces a Markov chain $\{g^t : t \in \mathbb{N} \cup \{0\}\}$ on the set of netwotks $\mathcal{G}$. The $2^{N(N-1)} \times 2^{N(N-1)}$ transition matrix $\Pi(x)$ has entries $\Pi(x)_{gw}$, $g, w \in \mathcal{G}$, which prescribe the probability of transitioning to w given the current period network g.

For each $g \in \mathcal{G}$, define $N(g) \coloneqq \{w \in \mathcal{G} \setminus \{g\} : \exists!\ (i, j) \in \mathcal{M}, g_{ij} \neq w_{ij}\}$ as the set of networks that differ from g in *exactly* one edge. Entries of $\Pi(x)$ take the form:

discrete choice and games (Aguirregabiria and Mira, 2010), though it is not innocuous. In our setting, it precludes randomness in the marginal effect of network covariates on utility (homophily in preferences), as taste shocks act as pure location shifts.

[8]Further notice that, differently from Mele (2017), we do *not* assume utilities admit a potential function. See Section 3.3 for a discussion.

[9]We denote by $\Delta(\mathcal{G})$ the set of all probability distributions on $\mathcal{G}$.

$$\Pi(x)_{gw} = \begin{cases} \rho((i,j),g,x)F_{\epsilon(g_{ij})-\epsilon(w_{ij})}(u_i(w,x)-u_i(g,x)) & \text{if } w \in N(g),\ g_{ij} \neq w_{ij} \\ \sum_{(i,j)\in\mathcal{M}} \rho((i,j),g,x)F_{\epsilon(1-g_{ij})-\epsilon(g_{ij})}\left(u_i(g,x)-u_i([1-g_{ij},g_{-ij}],x)\right) & \text{if } g = w \\ 0 & \text{elsewhere} \end{cases} \tag{2}$$

where $F_{\epsilon(1)-\epsilon(0)}$ and $F_{\epsilon(0)-\epsilon(1)}$ denote the distribution function of the difference in shocks.

Remark 2.1. The transition matrix is irreducible and aperiodic. Indeed, by Assumption 2.1 and Assumption 2.2, the first and second cases in (2) are always positive for any $g \in \mathcal{G}$. We can thus always reach any other network w starting from any g with positive probability in finite time (irreducibility). Since the chain is irreducible and contains a self-loop ($\Pi(x)_{gg} > 0$), it is also aperiodic.

We next look for (conditional) stationary distributions. Recall that a stationary distribution is an element $\pi(x) \in \Delta(\mathcal{G})$ satisfying $\pi(x) = \Pi(x)'\pi(x)$.

Remark 2.2. The transition matrix $\Pi(x)$ admits a unique stationary distribution, which is a direct consequence of the Perron-Froebenius theorem for nonnegative irreducible matrices (Horn and Johnson, 2012, Theorem 8.4.4). Moreover, as the chain is irreducible and aperiodic, we have that, for any $\pi_0 \in \Delta(\mathcal{G})$, $\lim_{t\to\infty}(\Pi(x)^t)'\pi_0 = \pi(x)$ (Norris, 1997, Theorem 1.8.3), so we may interpret the invariant distribution as a "long-run" distribution.

Remark 2.3. Note that the stationary distribution puts positive mass over all network configurations. Indeed, first note that there exists some $g_0 \in \mathcal{G}$ such that $\pi(g_0|x) > 0$. Fix $w \in \mathcal{G}$. Since the chain is irreducible, there exists $k \in \mathbb{N}$ such that $(\Pi(x)^k)_{g_0,w} > 0$, where $(\Pi(x)^k)_{g_0,w} > 0$ denotes the (g_0, w) entry of $\Pi(x)^k$. But $\pi(x) = (\Pi(x)^k)'\pi(x) \implies \pi(w|x) > 0$.

3 Identification

In this section, we study identification under many-network asymptotics. In particular, we assume we have access to a random sample (iid across c) of C networks $\{G_c^{T_0}, G_c^{T_1}, X_c\}_{c=1}^{C}$

stemming from the network formation game described in Section 2[10]. In this context, $G_c^{T_0}$ and $G_c^{T_1}$ are observations of network c over two (possibly nonconsecutive) periods[11] (labelled *first* and *second*); and X_c is the set of covariates associated with network c. We recall the law of X_c is $\mathbb{P}_X$, as X_c is a copy of X (i.e. a random variable with the same law as X). As in the previous section, realisations of X are denoted by small case letters, i.e. elements $x \in \mathcal{X}$.

Denote by $\Pi(X;\theta_0)$ the transition matrix under covariates X; and where $\theta_0 := ((u_i)_{i=1}^N, \rho)$ are the "true" parameters (functions). Further write τ_0 for the number of *rounds* of the network formation game taken place between the first and second period. Notice that we are able to identify $\Pi(X;\theta_0)^{\tau_0}$, the transition matrix to the power of the number of rounds of the network formation game which took place between the first and second period (τ_0), provided that the first-period conditional distribution, which we denote by $\pi_0(X)$, is such that $\pi_0(X) >> 0$ $\mathbb{P}_X$-a.s. To see this more clearly, suppose X were empty. In this case, we could consistently estimate $(\Pi^{\tau_0})_{gw}$, $g, w \in \mathcal{G}$, by $(\widehat{\Pi^{\tau_0}})_{gw} = \sum_{c=1}^C \mathbb{1}\{G_c^{T_0} = g, G_c^{T_1} = w\} / \sum_{c=1}^C \mathbb{1}\{G_c^{T_0} = g\}$[12], provided $\mathbb{P}[G_c^{T_0} = g] > 0$[13]. The next assumption summarises this requirement.

Assumption 3.1 (Full support). $\pi_0(g|X) > 0$ for all $g \in \mathcal{G}$ $\mathbb{P}_X$-a.s.

Since $\Pi(X;\theta_0)^{\tau_0}$ is identified from the data under Assumption 3.1, the identification problem subsumes to (denoting by Θ the parameter space[14]):

---

[10]Under large-network asymptotics, we would have access to a single (a few) network(s) with a large number of players.

[11]In Appendix B, we briefly discuss (non)identification when only one period of data stemming from the stationary distribution of the network formation game is available.

[12]In our setting, an agent is described by her set of exogenous characteristics W_i. Agents in different networks are "equal" if their covariates agree. Two networks are thus deemed equal if their adjacency matrices are equal up to symmetric permutations in their rows and columns; and if agents "agree" in both networks.

[13]The case where X has discrete support is similar to the case where no covariates are included. In the case where X contains continuous covariates, a consistent estimator would be given by a kernel estimator (Li and Racine, 2006). Clearly, these estimators would behave poorly in many practical settings (especially if the number of players is moderate to large), though we are not suggesting to use them in practice – we rely on consistency only as an indirect argument of identification.

[14]Abstracting from measurability concerns, this is the set of utilities and matching functions satifying

$$\forall(\theta,\tau),(\tilde{\theta},\tilde{\tau})\in\Theta\times\mathbb{N},\quad (\theta,\tau)\neq(\tilde{\theta},\tilde{\tau}) \implies \Pi(X;\theta)^{\tau}\neq\Pi(X;\tilde{\theta})^{\tilde{\tau}}$$

where the RHS inequality must hold with positive probability over the distribution of X (Newey and McFadden, 1994).

Observe that, in our statement of the identification problem, the number of rounds in the network formation game is assumed to be unknown. There is no reason to expect τ_0, the true number of rounds, to be known *a priori* by the researcher, unless the network formation algorithm has a clear empirical interpretation. Nonetheless, it is still possible to identify τ_0 under some assumptions. Observe that, for all $\theta\in\Theta$ and $x\in\mathcal{X}$, $\Pi(x;\theta)$ is irreducible and has strictly positive main diagonal. It is then easy to see that the number of strictly positive entries in $\Pi(x;\theta)^{\tau}$ is nondecreasing in τ. Moreover, this number is strictly increasing for $\tau\leq N(N-1)$[15] and does not depend on the choice of (x,θ). Thus, provided that $\tau_0\leq N(N-1)$, we can identify τ_0 by "counting"[16] the number of positive entries in $\Pi(X;\theta_0)^{\tau_0}$.

Assumption 3.2 (Upper bound on τ_0)**.** The number of rounds which took place in the network formation game between the first and second period (τ_0) is smaller than or equal to $N(N-1)$.

Remark 3.1. Under 3.2, τ_0 is identified (in $\{1,2\dots N(N-1)\}$).

A similar assumption can be found in Christakis et al. (2010), where it is assumed that $\tau_0=N(N-1)/2$ (it is an undirected network) and all meeting opportunities are played (though in an unknown order). In their setting, however, the assumption is mainly made in order to reduce the computational toll of evaluating the model likelihood (see Section 4.1.2 for a similar discussion); whereas in our environment we require it in identification. We also emphasise that the bound in assumption 3.2 will be more or less restrictive depending on the setting – knowledge of the particular application in mind should help to assess its appropriability. Finally, it should be noted that our main identification result (Proposition 3.1) holds irrespective of the bound, provided that τ_0 is identified or known *a priori*.

the assumptions in Section 2.

[15] $N(N-1)$ is the minimum number of rounds required for the probability of transitioning from a "fully empty" network to a "fully connected" network to be strictly positive.

[16] We provide a consistent estimator for τ_0 in Section 4.1.1.

Under assumption 3.2, we may thus assume, without loss, τ_0 is known.

In the next subsections, we discuss identification of θ.

3.1 Identification without restrictions

To illustrate the difficulty of identification without imposing further restrictions, let us briefly analyse identification of θ from $\Pi(X;\theta)$. Observe that identification of θ from $\Pi(X;\theta)$ is a necessary condition for identification of θ from $\Pi(X;\theta)^{\tau_0}$. Indeed, knowledge of $\Pi(X;\theta)$ implies knowledge of $\Pi(X;\theta)^{\tau_0}$. Thus, in a sense, our analysis in this subsection provides a "best-case" scenario for achieving identification without additional restrictions.

As we do not impose further restrictions in the model, we essentially view X as non-stochastic throughout the remainder of this subsection and suppress dependence of $\Pi(X;\theta)$ on X by writing $\Pi(\theta)$. In order to make the identification problem clearer, define, for all $g \in \mathcal{G}$, $w \in N(g)$ with $g_{ij} \neq w_{ij}$, $F_{ij}(g,w) := F_{\epsilon(g_{ij})-\epsilon(w_{ij})}(u_i(w,X)-u_i(g,X))$. Observe that $F_{ij}(g,w) + F_{ij}(w,g) = 1$. Write $\rho_{ij}(g)$ for $\rho((i,j),g,X)$. Observe that $\sum_{(i,j)\in\mathcal{M}} \rho_{ij}(g) = 1$. Let $\gamma := ((\rho_{ij}(g))_{g\in\mathcal{G},(i,j)\in\mathcal{M}}, \quad (F_{ij}(g,w))_{g\in\mathcal{G},w\in N(g),g_{ij}\neq w_{ij}})$ be a parameter vector, and γ_0 the "true" parameter. Observe that, under assumptions 2.1 and 2.2, the parameter space, which we denote by Γ, is a subset of $\mathbb{R}_{++}^{\dim\gamma}$, an open set. Put another way, $\Gamma = \{\gamma \in \mathbb{R}_{++}^{\dim\gamma} : F_{ij}(g,w) + F_{ij}(w,g) = 1, \sum_{(k,l)\in\mathcal{M}} \rho_{k,l}(g) = 1 \text{ for all } g \in \mathcal{G}, w \in N(g) \text{ with } g_{ij} \neq w_{ij}\}$. Identification from the transition matrix thus requires us to show that, for all $\gamma, \gamma' \in \Gamma$, $\gamma \neq \gamma' \implies \Pi(\gamma) \neq \Pi(\gamma')$, where $\Pi(\gamma)$ is the matrix in (2) constructed under γ. If we can uniquely recover γ from Π, then we can recover differences in utilities, $u_i(g,X) - u_i(w,X)$, for all $i \in \mathcal{I}$ and $g, w \in \mathcal{G}$, as $F_{\epsilon(1)-\epsilon(0)}$ is invertible under assumption 2.2. Levels (and thus θ) are then identified under a location normalisation on pay-offs (e.g. $u_i(g_0,X) = 0$ for all i and some g_0).

Observe that $\dim\gamma = N(N-1)2^{N(N-1)} + N(N-1)2^{N(N-1)}$, where the first summand is the dimension of $(p_{ij}(g) : g \in \mathcal{G}, (i,j) \in \mathcal{M})$ and the second term is the dimension of $(F_{ij}(g,w) : (i,j) \in \mathcal{M}, g \in \mathcal{G}, w \in N(g), w_{ij} \neq g_{ij})$. Matrix $\Pi(\gamma)$ has $2^{N(N-1)}(N(N-1)+1)$ strictly positive entries. The parameter space imposes $2^{N(N-1)}$ restrictions of the type $\sum_{(i,j)\in\mathcal{M}} \rho_{i,j}(g) = 1$ and $N(N-1)2^{N(N-1)}/2$ restrictions of the type $F_{ij}(g,w)+F_{ij}(w,g) = 1$. A simple order condition would thus require:

$$2^{N(N-1)}\left[2N(N-1)\right] \leq 2^{N(N-1)}\left[N(N-1)+1+N(N-1)/2\right] \implies N(N-1) \leq 2 \implies$$
$$\implies N \leq 2$$

and the model would be identified provided that $N = 2$. The point is that the map $\gamma \mapsto \Pi(\gamma)$ is nonlinear, so the order condition is nor necessary nor sufficient. Nonetheless, we are able to show directly that the model is identified when $N = 2$.

Claim 3.1. *If $N = 2$, then γ is identified from $\Pi(\gamma)$ under assumptions 2.2, 2.1 and 3.1.*

Proof. See Appendix D. □

Extending such a direct argument to $N > 2$ is not feasible, as Π is a $2^{N(N-1)} \times 2^{N(N-1)}$ matrix. Notice that the rank condition in Rothenberg (1971) for local identification is not satisfied (recall the previous order condition). The problem is that, for this condition to be sufficient for local nonidentification, the Jacobian of $\Pi(\gamma)$ must be rank-regular (i.e. it must have constant rank in a neighbourhood of γ) , which is not trivial to show. Of course, if that were the case, then we would know the model is nonidentified for $N > 2$.

Given the difficulty of establishing identification without imposing further restrictions even when $\Pi(X;\theta)$ is known (or $\tau_0 = 1$), in the next subsections we explore the identifying power of restrictions on: (i) how covariates affect utilities and the matching function; and (ii) how the network structure affects pay-offs and meetings. To make both the exposition and proofs clearer, in what follows we maintain the notation introduced in this section, and dependence of objects on covariates will remain implicit where there is no confusion.

3.2 Identification with covariates

In this subsection, we explore the identifying power of restrictions on covariates. We work in the environment where we observe two periods of data and τ_0 is assumed to be identified or known *a priori*. We will follow an identification at infinity approach in order to identify θ (Tamer, 2003; Bajari et al., 2010).

In order to make explicit the dependency in covariates, we write $X_i^u(g)$ for the covariates that enter the utility of agent i under network g, i.e. we shall write $u_i(g, X) = u_i(g, X_i^u(g))$

for all $i \in \mathcal{I}$, $g \in \mathcal{G}$. We use $X^m(g)$ for the covariates that enter the matching function under network g, i.e. $\rho((i,j),g,X) = \rho((i,j),g,X^m(g))$ for all $(i,j) \in \mathcal{M}$ and $g \in \mathcal{G}$. We also define $u_i(w, X_i^u(w)) - u_i(g, X_i^u(g)) =: \delta_{ij}(g, w, X_{ij}^u(g,w))$, the gain in utility from each choice, for all $g \in \mathcal{G}$, $w \in N(g)$, $w_{ij} \neq g_{ij}$. In this case, $X_{ij}^u(g,w)$ is the subvector of $[X_i^u(g), X_i^u(w)]$ with the covariates relevant in the marginal gain of i moving from g to w. We write $\Pi(X)^{\tau_0} = \Pi(X;\theta_0)^{\tau_0}$ for the observed transition matrix. Notice that, in our case, $X = [(X_{ij}^u(g,w))_{g\in\mathcal{G}, w\in N(g), w_{ij}\neq g_{ij}}, (X^m(g))_{g\in\mathcal{G}}]$. Finally, we will use the notation $A \setminus B$ for the subvector of A such that, up to permutations, $A = [A \setminus B, B]$.

We next impose the following assumptions.

Assumption 3.3 (Location normalisation)**.** There exists some $g_0 \in \mathcal{G}$, $u_i(g_0, X_i^u(g_0)) = 0$ for all $i \in \mathcal{I}$.

Such a normalisation is required in order to identify utilities in levels.

Our exclusion restriction is as follows:

Assumption 3.4 (Large support exclusion restriction)**.** For all $g \in \mathcal{G}$, $w \in N(g)$, $g_{ij} \neq w_{ij}$, there exists a $m \times 1$ subvector $Z_{ij}^u(g,w)$ of $X_{ij}^u(g,w)$, i.e. $X_{ij}^u(g,w) = [Z_{ij}^u(g,w), \tilde{X}_{ij}^u(g,w)]$, such that no covariate in $Z_{ij}^u(g,w)$ is an element of $X^m(g)$. Moreover, $Z_{ij}^u(g,w)$ admits a conditional Lebesgue density $f(Z_{ij}^u(g,w)|\tilde{X}_{ij}^u(g,w), X^m(g))$ that is positive a.e. (for almost all realisations of $[\tilde{X}_{ij}^u(g,w), X^m(g)]$); and there exists $\vec{r} \in \mathbb{R}^m$ s.t. $\lim_{t\to\infty} \delta_{ij}(g, w, [Z_{ij}^u(g,w) = t\vec{r}, \tilde{X}_{ij}^u(g,w)]) = \infty$.

Assumption 3.4 requires that, for each $g \in \mathcal{G}$, large support covariates be included in the marginal gain of each agent's choice under g, but excluded from the matching function under g. These covariates should admit, with positive probability, sufficiently "high" realisations s.t. the (conditional on X) probability of an agent "accepting" a transition from g once selected by the matching process can be made arbitrarily close to unity.

In what follows, write $N^k(g)$, $k \in \mathbb{N}$ and $g \in \mathcal{G}$, for the set of networks that differ from g in *exactly* k edges. The previous restrictions imply the next result:

Lemma 3.1. *Under Assumptions 2.1, 2.2, 3.1, 3.3 and 3.4, θ_0 is identified when $\tau_0 = 1$.*

Proof. Starting from some $w \in N(g_0)$, $w_{ij} \neq g_{0ij}$, we can identify $\rho_{ij}(g_0, X^m(g)) = \lim_{t\to\infty}(\Pi(X \setminus Z_{ij}^u(g_0,w), Z_{ij}^u(g_0,w) = t\vec{r}_{g_0w}))_{g_0w}$, which is valid under the (conditional) large support assumption. We are then able to identify $u_i(w, X_i^u(w))$ thanks to the normalisation on $u_i(g_0, X_i^u(g_0))$. Proceeding in a similar fashion iteratively on $w' \in N^2(g_0), N^3(g_0)\dots$, we identify all objects. $\square$

A sufficient condition for identification of θ_0 for any τ_0 known or identified is provided in the corollary below.

Corollary 3.1. *If $\theta \mapsto \Pi(X;\theta)$ is a.s. diagonalisable with the appropriate eigenvalue signs (nonnegative if τ_0 even), then, under the assumptions in Lemma 3.1, θ_0 is identified for any τ_0 known or identified.*

More generally, conditions for uniqueness of a stochastic τ_0th root of a transition matrix are quite complicated. See Higham and Lin (2011) for examples and sufficient conditions.

Remark 3.2. It should also be clear that the result in Lemma 3.1 would similarly hold if the large suppport variable were included in the matching function (but not in utilities). This may be more appropriate in some applied settings.

When $\tau_0 \geq 2$ and we do not know if $\Pi(X;\theta)$ is "appropriately" diagonalisable, we need stronger exclusion restrictions. We state a sufficient version (for all τ_0 identified or krnown) of this assumption below.

Assumption 3.5. The exclusion restriction in 3.4 holds as: no covariate in $Z_{ij}^u(g,w)$ is included in $[X^m(g), X^m(w), (X_{kl}^u(g,[1-g_{kl}, g_{-kl}]), X_{kl}^u(w,[1-w_{kl}, w_{-kl}]))_{(k,l)\neq(i,j)}]$; with $f(Z_{ij}^u(g,w)|X \setminus Z_{ij}^u(g,w))$ positive a.e. (for almost all realisations of $[X \setminus Z_{ij}^u(g,w)]$).

This stronger exclusion restriction requires that the large support covariates included in agent i's marginal gain of transitioning from g to w be excluded not only from the matching function under g; but also from the matching function under w and from all other agents' marginal gain of transitioning from g or w.

We next show identification when $\tau_0 = 2$ to get an idea of how the general case would look like. As one will see, the exclusion restriction given by Assumption 3.5 could be relaxed in this case, though the latter would then be insufficient for identification when $\tau_0 > 2$.

Lemma 3.2. *Suppose Assumptions 2.1, 2.2, 3.1, 3.3 and 3.5 hold. Then, if $\tau_0 = 2$, the model is identified.*

Proof. First notice that $\Pi(X;\theta)^2$ takes the form:

$$(\Pi(X;\theta)^2)_{gw} = \begin{cases} \rho_{ij}(g)F_{ij}(g,[w_{ij},g])\rho_{kl}([w_{ij},g])F_{kl}([w_{ij},g],w)+ & \\ +\rho_{kl}(g)F_{kl}(g,[w_{kl},g])\rho_{ij}([w_{kl},g])F_{ij}([w_{kl},g],w) & \text{if } w \in N^2(g), w_{ij} \neq g_{ij}, w_{kl} \neq g_{kl} \\ \Pi_{gg}\rho_{ij}(g)F_{ij}(g,w) + \rho_{ij}(g)F_{ij}(g,w)\Pi_{ww} & \text{if } w \in N(g), w_{ij} \neq g_{ij} \\ \Pi_{gg}\Pi_{gg} + \sum_{s\in N(g)} \Pi_{gs}\Pi_{sg} & \text{if } g = w \\ 0 & \text{otherwise} \end{cases} \tag{3}$$

Fix $g \in \mathcal{G}$, $w \in N(g)$, $g_{ij} \neq w_{ij}$. Notice that, by driving $F_{ij}(g,w) \to 1$ and $F_{pq}(g,m) \to 0$ for all $m \in N(g) \setminus \{w\}$, $g_{pq} \neq m_{pq}$, the term $(\Pi^2)_{gg}$ identifies:

$$\lim_{t^*}(\Pi^2)_{gg} = (1-\rho_{ij}(g))^2$$

where $\lim_{t^*}$ is shorthand for the appropriate limit[17]. Since $\lim_{t^*}(\Pi^2)_{gg} \in [0,1]$, we can uniquely solve for $\rho_{ij}(g)$, thus establishing identification of $\rho_{ij}(g)$.

Next, observe that, by taking $F_{pq}(g,m) \to 0$ for all $m \in N(g) \setminus \{w\}$, $F_{pq}(w,m) \to 0$ for all $m \in N(w) \setminus \{g\}$, the term Π_{gw} identifies:

$$\begin{aligned} \lim_{t^{**}} \Pi_{gw} = (1-\rho_{ij}(g)F_{ij}(g,w))\rho_{ij}(g)F_{ij}(g,w) + \rho_{ij}(g)F_{ij}(g,w)(1-\rho_{ij}(w)F_{ij}(w,g)) = \\ \rho_{ij}(g)F_{ij}(g,w)(2-\rho_{ij}(w) + (\rho_{ij}(w)-\rho_{ij}(g))F_{ij}(g,w)) \end{aligned}$$

where $\lim_{t^{**}}$ is shorthand for the appropriate limit[18]. Since the right-hand term is strictly increasing in $F_{ij}(g,w)$, the "true" $F_{ij}(g,w)$ uniquely solves the equation, thus establishing identification. □

The previous argument suggests a procedure for the general case.

Proposition 3.1. *Suppose Assumptions 2.1, 2.2, 3.1, 3.3 and 3.5 hold. Then the model is identified for any τ_0 known or identified.*

Proof. Fix $g \in \mathcal{G}$, $w \in N(g)$, $g_{ij} \neq w_{ij}$. Observe that:

$$(\Pi^{\tau_0})_{gg} = \sum_{m\in N(g)\cup\{g\}} (\Pi^{\tau_0-1})_{gm}\Pi_{mg}$$

We first prove the following claim:

[17] I.e. a limit that drives $F_{ij}(g,w) \to 1$ and $F_{pq}(g,m) \to 0$ for all $m \in N(g) \setminus \{w\}$, $g_{pq} \neq m_{pq}$.

[18] I.e. a limit that drives $F_{pq}(g,m) \to 0$ for all $m \in N(g) \setminus \{w\}$, $F_{pq}(w,m) \to 0$ for all $m \in N(w) \setminus \{g\}$.

Claim. *Under a limit which drives* $F_{pq}(w,m) \to 0$ *for all* $m \in N(w)$, $m_{pq} \neq w_{pq}$, *and* $F_{kl}(g,s) \to 0$ *for all* $s \in N(g) \setminus \{w\}$, $s_{kl} \neq g_{kl}$, *we have* $\lim_{t^*}(\Pi^{\tau_0})_{gg} = (1-\rho_{ij}(g))^{\tau_0}$, *where* $\lim_{t^*}$ *is shorthand for the appropriate limit.*

Proof. The case $\tau_0 = 1$ is readily verified by driving $F_{kl}(g,s) \to 0$ for all $s \in N(g) \setminus \{w\}$, $s_{kl} \neq g_{kl}$ and $F_{ij}(g,w) \to 1$. For $\tau_0 > 1$, we begin by noticing that we may drive $(\Pi^{\tau_0-1})_{gm} \to 0$ for all $m \in N(g) \setminus \{w\}$. Since m and g differ in exactly one edge (say, $m_{pq} \neq g_{pq}$), a transition in edge pq must appear in every summand in $(\Pi^{\tau_0-1})_{gm}$. Indeed, $(\Pi^{\tau_0-1})_{gm}$ sums over all possible transitions in edge pq from value g_{pq} to m_{pq} in $\tau_0 - 1$ rounds. Put another way, for each summand in $(\Pi^{\tau_0-1})_{gm}$, there exists $t \in \{0, 1 \ldots \tau_0 - 2\}$, $g^t_{pq} = g_{pq}$ and $g^{t+1}_{pq} = m_{pq}$[19]. Fix a summand in $(\Pi^{\tau_0-1})_{gm}$. We analyse the following cases:

1. There exists $t \in \{0, 1 \ldots \tau_0 - 2\}$, $g^t_{pq} = g_{pq}$, $g^{t+1}_{pq} = m_{pq}$ **and** $g^t = g$. In this case, by taking $F_{pq}(g,m) \to 0$, we drive the summand to 0.

2. For all $t \in \{0, 1 \ldots \tau_0 - 2\}$ such that $g^t_{pq} = g_{pq}$, $g^{t+1}_{pq} = m_{pq}$, we have $g^t \neq g$. Take t^* to be the smallest t satisfying the above. Observe that $t^* > 0$, as $g^0 = g$ (we always start at g). Since $g^{t^*} \neq g$, there exists $t' < t^*$, $g^{t'} = g$ and $g^{t'+1} = z$, $z \in N(g)$. If there exists some t' satisfying this property such that $z \neq w$ (with $g_{kl} \neq z_{kl}$), then driving $F_{kl}(g,z) \to 0$ vanishes the term. If, for all such t', $z = w$, take t^{**} to be the maximum of such t'. Observe that $t^{**} < t^*$. If $g^{t^*} = w$, we may safely drive $F_{pq}(w, [m_{pq}, w_{-pq}]) \to 0$. If not, then $t^{**} + 1 < t^*$ and there exists a transition from w to some element in $N(w)$ which we can safely drive to 0.

Since the above argument holds irrespective of the summand (the common limit will vanish all terms), we conclude $(\Pi^{\tau_0-1})_{gm} \to 0$. Since $F_{ij}(g,w) \to 1$, $\lim_{t^*}(\Pi^{\tau_0-1})_{gw}\Pi_{wg} = 0$[20]. The common limit in the statement of the claim thus leaves us with:

$$\lim_{t^*}(\Pi^{\tau_0})_{gg} = \lim_{t^*}(\Pi^{\tau_0-1})_{gg} \lim_{t^*}(\Pi)_{gg} = \lim_{t^*}(\Pi^{\tau_0-1})_{gg}(1-\rho_{ij}(g))$$

Induction then yields the desired result. $\square$

[19]Recall g^t is the stochastic process on $\mathcal{G}$ induced by the game.

[20]In all previous arguments, we implicitly use the sandwich lemma to infer that, if one term of the product goes to zero, the whole product does. This is immediate as we're working with products of probabilities.

Since $\lim_{t^*}(\Pi^{\tau_0})_{gg} \in [0,1]$, we can uniquely solve for $\rho_{ij}(g)$, thus establishing identification.

Next, we proceed to identification of $F_{ij}(g,w)$. Note that:

$$(\Pi^{\tau_0})_{gw} = \sum_{m \in N(w)\cup\{w\}} (\Pi^{\tau_0-1})_{gm}\Pi_{mw}$$

We then prove the following claim:

Claim. *Under a limit which drives $F_{pq}(w,m) \to 0$ for all $m \in N(w)\setminus\{g\}$, $m_{pq} \neq w_{pq}$, and $F_{kl}(g,s) \to 0$ for all $s \in N(g)\setminus\{w\}$, $s_{kl} \neq g_{kl}$:*

$$\lim_{t^{**}}(\Pi^{\tau_0})_{gw} = \lim_{t^{**}}(\Pi^{\tau_0-1})_{gg}\rho_{ij}(g)F_{ij}(g,w) + \lim_{t^{**}}(\Pi^{\tau_0-1})_{gw}(1-\rho_{ij}(w)F_{ij}(w,g))$$

where $\lim_{t^{**}}$ is shorthand for the appropriate limit. We also have that, under such a limit:

$$\lim_{t^{**}}(\Pi^{\tau_0})_{gw} + \lim_{t^{**}}(\Pi^{\tau_0})_{gg} = 1$$

Proof. Notice that, for $\tau_0 = 1$, we have:

$$\lim_{t^{**}}(\Pi)_{gw} = \rho_{ij}(g)F_{ij}(g,w)$$
$$\lim_{t^{**}}(\Pi)_{gg} = (1-\rho_{ij}(g)F_{ij}(g,w))$$

These expressions follow directly from the limit being taken and equation (2).

Consider next the case $\tau_0 > 1$. Observe that the limit in the statement of the lemma drives $(\Pi^{\tau_0-1})_{gm} \to 0$ for all $m \in N(w)\setminus\{g\}$. Indeed, notice that, if $m \in N(w)\setminus\{g\}$, then $m \in N^2(g)$. Recall $(\Pi^{\tau_0-1})_{gm}$ sums over all possible transitions from g to m in $\tau_0 - 1$ rounds. Fix a summand in $(\Pi^{\tau_0-1})_{gm}$. If a transition from g occurs at pair $(a,b) \in \mathcal{M}$, $(a,b) \neq (i,j)$, then the limit vanishes the term. If all transitions from g occur at pair (i,j) (i.e. g only transitions to w), a transition from w must occur, since $m \in N^2(g)$. If w only transitions to g, then either $m = g$ or $m = w$, which is not true. Therefore, there exists a transition from w to some $z \in N(w)\setminus\{g\}$, so we can vanish the summand. The limit in the statement thus drives the term $(\Pi^{\tau_0-1})_{gm}$ to zero.

From the above discussion, we thus get:

$$\lim_{t^{**}}(\Pi^{\tau_0})_{gw} = \lim_{t^{**}}(\Pi^{\tau_0-1})_{gg}\Pi_{gw} + \lim_{t^{**}}(\Pi^{\tau_0-1})_{gw}\Pi_{ww} =$$

$$\lim_{t^{**}}(\Pi^{\tau_0-1})_{gg}\rho_{ij}(g)F_{ij}(g,w) + \lim_{t^{**}}(\Pi^{\tau_0-1})_{gw}(1-\rho_{ij}(w)F_{ij}(w,g))$$

which establishes the first part of the claim.

Next, we notice that, under the limit in the statement of the claim:

$$\lim_{t^{**}}(\Pi^{\tau_0})_{gg} = \lim_{t^{**}}(\Pi^{\tau_0-1})_{gg}(1-\rho_{ij}(g)F_{ij}(g,w)) + \lim_{t^{**}}(\Pi^{\tau_0-1})_{gw}\rho_{ij}(w)F_{ij}(w,g)$$

This follows from observation that, in the proof of the *previous* claim, we can still drive $(\Pi^{\tau_0-1})_{gm} \to 0$ for all $m \in N(g) \setminus \{w\}$ even though $F_{ij}(g,w)$ does not vanish[21]. We are thus left with the terms related to staying in g or transitioning to w in $\tau_0 - 1$ rounds.

Finally, the second part of the claim can be asserted by noticing that $\lim_{t^{**}}(\Pi)_{gw} + \lim_{t^{**}}(\Pi)_{gg} = 1$ and applying this fact inductively on the expression for $\lim_{t^{**}}(\Pi^{\tau_0})_{gw} + \lim_{t^{**}}(\Pi^{\tau_0})_{gg} = \lim_{t^{**}}(\Pi^{\tau_0-1})_{gw} + \lim_{t^{**}}(\Pi^{\tau_0-1})_{gg}$. □

To establish identification of $F_{ij}(g,w)$, we need to show that the the expression for $\lim_{t^{**}}(\Pi^{\tau_0})_{gw}$ is strictly increasing in $(0,1)$ as a function of $F_{ij}(g,w)$. Denoting by $D_{F_{ij}(g,w)} \lim_{t^{**}}(\Pi^{\tau_0})_{gw}$ the derivative of $\lim_{t^{**}}(\Pi^{\tau_0})_{gw}$ as a function of $F_{ij}(g,w)$, we get:

$$\begin{aligned} D_{F_{ij}(g,w)} \lim_{t^{**}}(\Pi^{\tau_0})_{gw} = D_{F_{ij}(g,w)} \lim_{t^{**}}(\Pi^{\tau_0-1})_{gw}(1-\rho_{ij}(w) + (\rho_{ij}(w)-\rho_{ij}(g))F_{ij}(g,w)) \\ +\rho_{ij}(g)(1-\lim_{t^{**}}(\Pi^{\tau_0-1})_{gw}) + \rho_{ij}(w)\lim_{t^{**}}(\Pi^{\tau_0-1})_{gw} \end{aligned}$$

where we used that $\lim_{t^{**}}(\Pi^{\tau_0})_{gw} + \lim_{t^{**}}(\Pi^{\tau_0})_{gg} = 1$ and $F_{ij}(g,w) + F_{ij}(w,g) = 1$. By noticing $\rho_{ij}(g)(1-\lim_{t^{**}}(\Pi^{\tau_0-1})_{gw}) + \rho_{ij}(w)\lim_{t^{**}}(\Pi^{\tau_0-1})_{gw} \geq \min\{\rho_{ij}(g), \rho_{ij}(w)\} > 0$ and applying induction on the fact that $D_{F_{ij}(g,w)} \lim_{t^{**}}(\Pi)_{gw} \geq 0$, we get that the derivative is strictly positive in $(0,1)$, thus showing the map is invertible and establishing identification of $F_{ij}(g,w)$. □

In Appendix E, we show that identification would similarly hold if an analagous exclusion restriction on the matching function were true.

[21]Suppose g only transitions to w. Since $m \neq w$, w must transition to some other $z \in N(w)$. If w only transitions to g, then either $m = g$ or $m = w$, which is not true. Therefore, we can always vanish a summand in $(\Pi^{\tau_0-1})_{gm}$, even though $F_{ij}(g,w)$ does not vanish.

Remark 3.3 (A possible parametrisation)**.** If we take $\rho((i,j), X^m(g)) = \frac{\exp(\alpha_g' X_{ij}^m(g))}{\sum_{(k,l)\in\mathcal{M}} \exp(\alpha_g' X_{kl}^m(g))}$ and $u_i(g, X_i^u(g)) = \beta_g' X_i^u(g)$, where $X_i^u(g)$ may include other individuals' characteristics, then the model is identified under the previous restrictions, provided the usual rank conditions hold (cf. Amemiya (1985, p.286-292); also McFadden (1973))[22].

3.3 Identification via the network structure

In Mele (2017), it is assumed that $\rho_{ij}(g, X) = \rho_{ij}([1-g_{ij}, g_{-ij}], X)$ for all $g \in \mathcal{G}$, i.e. meeting probabilities do not depend on the presence of a link between ij. Notice that this constitutes an exclusion restriction with identifying power in our environment. Indeed, if $\tau_0 = 1$ and this hypothesis holds, $\Pi_{g,[1-g_{ij},g_{-ij}]}/\Pi_{[1-g_{ij},g_{-ij}],g} = F_{ij}(g, [1-g_{ij}, g_{-ij}])/(1-F_{ij}(g, [1-g_{ij}, g_{-ij}]))$, which establishes identification of utilities (and meeting probabilities thereupon) under a location normalisation. For $\tau_0 > 1$, identification is not that immediate[23], but we can rely on a sufficient condition such as "appropriate" diagonalisability as in Corollary 3.1 to achieve identification.

If we further assume that: (1) taste shocks are independent EV type 1; and (2) utility functions admit a potential function[24] $Q : \mathcal{G} \times \mathcal{X} \mapsto \mathbb{R}$; then the model's (conditional on X) network stationary distribution is in the exponential family, i.e. $\pi(g|X) \propto \exp(Q(g, X))$ (Mele, 2017, also Appendix C). If we assume that $\{G_c^{T_0}\}_c$ are drawn from the model's stationary distribution, then the model's potential (and hence marginal utilities) is identified under standard assumptions (Newey and McFadden, 1994). Provided that utility functions in the game played before period T_0 remain unaltered in the game played between period T_0 and period T_1, we can use the period T_0 distribution to help identify the model[25]. A

[22]Specifically. we would require $\mathbb{E}[(X_i^u(w) - X_i^u(g))(X_i^u(w) - X_i^u(g))']$ to have full rank for all $i \in \mathcal{I}$, $g \in \mathcal{G}$, $w \in N(g)$, with $X_i^u(g_0) = \vec{0}$ for all $i \in \mathcal{I}$; and $\mathbb{E}[\tilde{X}^m(g)'\tilde{X}^m(g)]$ to have full rank for all g, where $\tilde{X}^m(g)' = \begin{bmatrix}(X_{1,1}^m(g) - \bar{X}^m(g))' & (X_{1,2}^m(g) - \bar{X}^m(g))' & \ldots & (X_{N,(N-1)}^m(g) - \bar{X}^m(g))'\end{bmatrix}$ and $\bar{X}^m(g) = \sum_{(i,j)\in\mathcal{M}} \rho((i,j), g, X^m(g)) X_{ij}^m(g)$.

[23]Notice from (3), nonetheless, that $(\Pi^2)_{g,[1-g_{ij},g_{-ij}]}/(\Pi^2)_{[1-g_{ij},g_{-ij}],g} = F_{ij}(g, [1-g_{ij}, g_{-ij}])/(1-F_{ij}(g, [1-g_{ij}, g_{-ij}]))$, which identifies utilities. Identification of meeting probabilities is not immediate in this case, though.

[24]A *potential function* is a map $Q : \mathcal{G} \times \mathcal{X} \mapsto \mathbb{R}$ satisfying $Q([1, g_{-ij}], X) - Q([0, g_{-ij}, X]) = u_i([1, g_{-ij}], X) - u_i([0, g_{-ij}], X)$ for all $(i,j) \in \mathcal{M}$, $g \in \mathcal{G}$.

[25]Observe we do not need to assume matching functions remain unaltered, just that before period T_0

necessary condition for this equality in utilities, provided that between period T_0 and period T_1 the matching function satisfies the restriction in Mele (2017), is that the period T_0 network distribution equals the period T_1 distribution[26]. This is a testable assumption.

More generally, we could try to achieve identification by restricting how pay-offs are affected by the network structure. This approach is followed by de Paula et al. (2018) and Sheng (2014), where it is assumed that network observations are pairwise-stable realisations of a (static) simultaneous-move complete information game. In their setting, pairwise stability only enables partial identification. We recognise that further restrictions on how the network structure affects pay-offs may enable point-identification in our setting, though we do not try to analyse these conditions in a general environment.

4 Estimation

In this section, we will analyse estimation. We have access to a sample of C networks, $\{G_c^{T_0}, G_c^{T_1}, X_c\}_{c=1}^{C}$, stemming from the network formation game previously described.

4.1 Frequentist estimation

4.1.1 Estimating τ_0

We first propose to estimate τ_0 as follows:

$$\hat{\tau} = \max_c \{ \| G_c^{T_1} - G_c^{T_0} \|_1 \} \tag{4}$$

where $\|\cdot\|_1$ is the L1 norm (a matrix viewed as its vectorisation). This estimator is intuitive: it amounts to "counting" in each network the number of differing edges between periods and then taking the maximum. It turns out that, under iid sampling and the bound in 3.2, $\hat{\tau} \overset{a.s.}{\to} \tau_0$.

Lemma 4.1. *Suppose $\{G_c^{T_0}, G_c^{T_1}, X_c\}_{c=1}^{C}$ is a random sample (iid across c). Under assumptions 2.1, 2.2 and 3.2, $\hat{\tau} \overset{a.s.}{\to} \tau_0$.*

they obeyed the restriction in Mele (2017).

[26]If the equality is conditional on X, this is also a sufficient condition, as it is assumed the potential before period T_0 is identified from $\pi_0(g|X)$.

Proof. Consider the event $\{\hat{\tau} \to \tau_0\}^{\complement}$. Observe that $\{\hat{\tau} \to \tau_0\}^{\complement} = \cap_{c\in\mathbb{N}}\{\|G_c^{T_1} - G_c^{T_0}\|_1 < \tau_0\}$. Fix $k \in \mathbb{N}$ and notice that:

$$\mathbb{P}\left[\cap_{c=1}^{k}\{\|G_c^{T_1} - G_c^{T_0}\|_1 < \tau_0\}\right] = \phi^k$$

where ϕ is the ex-ante (unconditional) probability that $G_c^{T_1}$ differs from $G_c^{T_0}$ in strictly less than τ_0 edges. Since $\mathbb{P}[\|G_c^{T_1} - G_c^{T_0}\|_1 = \tau_0] \in (0,1)$ (which follows from 2.1, 2.2 and 3.2), we have $\phi \in (0,1)$. Passing k to the limit and using continuity of $\mathbb{P}$ from above, we conclude the desired result. $\square$

We may extend the above result in order to allow for a sequence of *independent* observations stemming from a game with common parameters but allowing the distribution of covariates (and the number of players) to vary. In this case, we must restrict the distribution of covariates to not shift "too" much to high-probability regions. Formally, the proof would change as the ex-ante probability would now depend on c, i.e. we would have $\prod_{c=1}^{k}\phi_c$ in the formula. If $\limsup_{c\to\infty}\phi_c < 1$, we would get the same result.

4.1.2 Estimation of preference and meeting parameters

Let vector $\beta_0 \in \mathbb{B} \subseteq \mathbb{R}^l$ encompass a parametrisation of peferences and meetings, i.e. $u_i(g, X) = u_i(g, X; \beta_0)$ and $\rho_{ij}(g, X) = \rho_{ij}(g, X; \beta_0)$ for all $(i,j) \in \mathcal{M}$, $g \in \mathcal{G}$. The network log-likelihood, conditional on X_c, τ_0 and $G_c^{T_0}$ is:

$$l_c(G_c^{T_1}|G_c^{T_0}, X_c; \tau_0, \beta) = \sum_{g\in\mathcal{G}} \mathbb{1}\{G_c^{T_1} = g\} \ln\left((\Pi(X_c;\beta)^{\tau_0})_{G_c^{T_0} g}\right)$$

and the sample log-likelihood, under an independent sequence of observations, is:

$$\mathcal{L}(\{G_c^{T_1}\}_{c=1}^{C}|\{G_c^{T_0}\}_{c=1}^{C}; \tau_0, \beta) = \sum_{c=1}^{C}\sum_{g\in\mathcal{G}} \mathbb{1}\{G_c^{T_1} = g\} \ln\left((\Pi(X_c;\beta)^{\tau_0})_{G_c^{T_0} g}\right)$$

The second-step MLE estimator will thus be:

$$\hat{\beta}_{\text{MLE}} \in \operatorname{argmax}_{\beta\in\mathbb{B}} \mathcal{L}(\{G_c^{T_1}\}_{c=1}^{C}|\{G_c^{T_0}\}_{c=1}^{C}; \hat{\tau}, \beta)$$

where $\hat{\tau}$ is the estimator discussed in the previous section. Observe that this formulation can be easily modified to accommodate for observations of networks with different numbers of players, provided they have common parameters.

Numerically, computation of the likelihood is complicated by the fact we need to sum over all walks between $G_c^{T_0}$ and $G_c^{T_1}$. For small τ_0, this is feasible, but for higher values of τ_0, it gets impractical.

Given the above difficulty, an interesting alternative is to work with simulation-based methods which allow us to bypass direct evaluation of the model likelihood. A simulated method of moments estimator is a possibility in our case, though the nonsmoothness of the objective function (which would involve indicators of simulated network observations) as well as the poor properties of GMM estimators with many moment conditions (transition probabilities) in small samples (Newey and Smith, 2004) are somewhat unappealing. An indirect inference approach is also unappealing, as low-dimensional sufficient statistics are unknown in our context. Instead, we opt for a Bayesian approach, which we describe in detail next.

4.2 Bayesian estimation

As emphasised in Section 4.1.2, the difficulty of evaluating the model likelihood lies in computing all walks between $G_c^{T_0}$ and $G_c^{T_1}$. For a given $\tau \in \mathbb{N}$, there are $[N(N-1)+1]^\tau$ walks starting from $G_c^{T_0}$ and ending in some network $g \in \mathcal{G}$. Evaluating the model likelihood would require summing over all walks ending in $G_c^{T_1}$. A "recursive" approach for evaluating the model likelihood would consist in, for each $c \in \{1, 2 \dots C\}$, "writing down" the formula for each walk iteratively, i.e. starting from $G_c^{T_0}$, compute all possible N(N-1) transitions in the first round; then, for each of these N(N-1) possible transitions, compute the N(N-1) transitions in the second round and multiply each of these probabilities by the probability of the associated transition in the first round, and so on; and then summing over all walks ending in $G_c^{T_1}$. Walks that "strand off" from $G_c^{T_1}$ in some round $r < \tau$ can be excluded from the next steps in the recursion[27], which ameliorates the computational toll, but does not solve it.

An approach in Bayesian estimation that bypasses evaluating the model likelihood is likelihood-free estimation (Sisson and Fan, 2011), also known as Approximate Bayesian

[27]By "strand off" we mean a path of realisations of the stochastic process g_c^t up to round r such that the probability of reaching $G_c^{T_1}$ in $\tau_0 - r$ rounds is 0.

Computation (ABC)[28]. This method has a close correspondence with Nonparametric (Frequentist) Econometrics (Blum, 2010) and Indirect Inference (cf. Frazier et al. (2018) for a discussion). The methodology basically requires the researcher be able to draw a sample (or statistics thereof) from the model given the parameters. Algorithm 1 outlines the simplest accept-reject ABC algorithm in our setting, where S is the maximum number of iterations and $p_0(\beta, \tau)$ is a prior distribution over $\mathbb{B} \times \mathbb{N}$.

Algorithm 1 Basic Accept-Reject ABC algorithm

define some tolerance $\epsilon > 0$
define a vector of m statistics $T : \mathcal{G}^C \mapsto \mathbb{R}^m$
compute the observed sample statistics $T_{\text{obs}} \coloneqq T(\{G_c^{T_1}\}_{c=1}^C)$
for $s \in \{1, 2 \ldots S\}$ **do**
  draw $(\beta_s, \tau_s) \sim p_0$
  generate an artificial sample $\{\tilde{G}_c^{T_1}\}_{c=1}^C$ given $\{G_c^{T_0}, X_c\}_{c=1}^C$ and (β_s, τ_s)
  compute the simulated statistic $T_s \coloneqq T(\{\tilde{G}_c^{T_1}\})$
  if $\|T_s - T_{\text{obs}}\| \leq \epsilon$, accept (β_s, τ_s)
end for

In practice, a few improvements can be made upon Algorithm 1 (Li and Fearnhead, 2018). First, we can use importance sampling: instead of drawing from the prior, we may draw from a proposal distribution $q_0(\beta, \tau)$ such that supp $p_0 \subseteq$ supp q_0. Accepted draws should then be associated with weights $w_s \coloneqq p_0(\beta_s, \tau_s)/q_0(\beta_s, \tau_s)$. Second, we can use a "smooth" rejection rule, i.e. we accept a draw with probability $K(\|T_s - T_{\text{obs}}\|/\epsilon)$, where $K(\cdot)$ is a rescaled univariate kernel such that $K(0) = 1$.

In this methodology, there are two crucial choices to be made by the researcher. One is the tolerance parameter. Here, we can use the recommendations in Li and Fearnhead (2018): we may choose ϵ so the algorithm produces a "reasonable" acceptable rate. The

[28] An alternative and well-known approach to simplify, but not bypass, a complicated model likelihood is data-augmentation (Hobert, 2011). This alternative is not very useful in our setting, though, due to the dimensionality of the support of the meeting process. Thus, an approach that conditions the likelihood on the (unobserved) matching process (notice that this reduces the number of walks starting at $G_c^{T_0}$ from $[N(N-1)+1]^\tau$ to 2^τ) is not very useful in our case, since we still have to draw from the matching process distribution conditional on the data and the model parameters.

second important choice is the vector of statistics to be used. This is closely related to identification: for a proper working of the ABC algorithm, it is crucial the binding function (the map $b(\beta_s, \tau_s) := \text{plim}_{C\to\infty} T_s$) associated with the chosen vector of statistics identifies the model.

In Appendix F, we show in our setting that, by taking the vector of statistics to be the data, if we let $\epsilon \to 0$ as $S \to \infty$, then the mean of the accepted draws $h(\theta_s)$ converges in probability to the expectation of $h(\cdot)$ with respect to the posterior, where $h(\cdot)$ is a function with finite first moments (with respect to the prior). This result motivates computation of approximations to the posterior mean and credible intervals, which we undertake in our application.

5 Application

In our application, we consider data on friendship networks from Pinto and Ponczek (2017). The dataset comprises information on 3rd and 5th graders from 30 elementary schools in Recife, Brazil. Data on students' traits and intraclassroom friendship networks was collected at the beginning (baseline) and at the end (followup) of the 2014 school year[29]. Once missing observations are removed, our working sample comprises 161 classrooms (networks) totalling 1,589 students[30].

As a first step in our analysis, we attest that homophily is indeed a salient feature of our data. For that, we run regressions of the type:

$$g_{ij,c,1} = \beta' W_{ij,c,0} + \alpha_i + \gamma_j + \epsilon_{ij,c} \tag{5}$$

where $g_{ij,c,1}$ equals 1 if, in classroom c, individual i nominates j as a friend *at the followup period.* Vector $W_{ij,c,0}$ consists of pairwise distances in gender, age (in years) and measures of

[29]More specifically, students could nominate up to 8 classmates in each of the three categories: classmates whom they would (i) study with, (ii) talk to or (iii) play with. We consider an individual a friend if she falls in at least one of the three lists. No student exceeds 8 friends, as in most cases the three criteria coincide. In our raw dataset, only 1.01% of students reported 8 friends at the baseline, and this number falls to 0.3% at the followup.

[30]Figure 2 in the Appendix plots one such network.

cognitive and non-cognitive skills between i and j *at the baseline period*[31]. The specification controls for sender (α_i) and receiver (γ_j) fixed effects. We cluster standard errors at the classroom level.

Column 1 in Table 1 reports estimates obtained from running the above specification. Results indicate homophily is pervasive, e.g. same-sex classmates are, on average, 19.8 pp more likely to be friends, relatively to boy-and-girl pairs.

[31] Table 4 in the Appendix presents summary statistics of our dyad-level covariates.

Table 1: Dyadic regressions

	Dependent variable:	
	edge	
	(1)	(2)
distance in classlist		-0.002^{***}
		(0.0005)
distance in age	-0.018^{***}	0.002
	(0.006)	(0.003)
distance in gender	-0.198^{***}	-0.186^{***}
	(0.006)	(0.006)
distance in cognitive skills	-0.254^{***}	-0.127^{***}
	(0.049)	(0.034)
distance in conscientiousness	-0.031^{***}	-0.019^{***}
	(0.009)	(0.006)
distance in neuroticism	-0.017^{**}	-0.023^{***}
	(0.008)	(0.005)
Sender fixed effects?	Yes	No
Receiver fixed effects?	Yes	No
Time effect?	Yes	No
Observations	17,736	17,736
R^2	0.323	0.064
Adjusted R^2	0.185	0.063

Note: *p<0.1; **p<0.05; ***p<0.01

Standard errors clustered at the classroom-level in parentheses.

As discussed in Section 3, nonparametric identification of our model requires a large-support pair-level covariate which enters a pair's marginal utility, but is excluded from other pairs' marginal utilities and the matching function (in the current and adjacent networks). Alternatively, we require a large-support pair-level covariate that enters the matching function, but is excluded from pairs' marginal utilities. We focus on the latter case and propose to use the distance between two students in the alphabetically-ordered classlist, interacted with one minus an indicator of a link between the pair in the current network, as our "instrument". The intuition is that the position in the classlist should affect the odds of a pair meeting, e.g. through class activities in groups, but should not enter marginal utilities. We also speculate this mechanism is important for pairs that are not currently friends.

Column 2 provides some reduced-form evidence of relevance of our classlist distance variable. We run the specification in (5), but include our classlist distance variable and exclude sender and receiver fixed effects. The covariate is statistically significant at the 1% level and with the expected sign: all else equal, classmates "one more student away" in the classlist are 0.2 pp less likely to be friends.

In estimating our network formation model, we parameterise the matching function as follows. For individuals i and j in classroom c, we specify:

$$\rho((i,j), g, X_c) \propto \exp\left(\beta_m' W_{ij,c,0} + \delta_0 g_{ij} + \delta_1 (1 - g_{ij}) Z_{ij,c,0}\right)$$

where $W_{ij,c,0}$ and $Z_{ij,c,0}$ denote respectively our vector of pair-level covariates and classlist distance variable *at the baseline*. Notice our specification allows for explicit dependence of meetings on a previous-period link between agents. The utility of individual i in classroom c follows a linear parameterisation of Mele (2017), i.e.

$$\begin{aligned} u_i(g, X_c) = & \underbrace{\sum_{k \neq i} \beta_{ud}' \begin{pmatrix} 1 \\ W_{ik,c,0} \end{pmatrix} g_{ik}}_{\text{direct links}} + \underbrace{\sum_{k \neq i} \beta_{ur}' \begin{pmatrix} 1 \\ W_{ik,c,0} \end{pmatrix} g_{ki}}_{\text{mutual links}} + \\ & + \underbrace{\sum_{k \neq i} g_{ik} \sum_{\substack{l \neq i \\ l \neq k}} \beta_{un}' \begin{pmatrix} 1 \\ W_{il,c,0} \end{pmatrix} g_{kl}}_{\text{indirect links}} + \underbrace{\sum_{k \neq i} g_{ik} \sum_{\substack{l \neq i \\ l \neq k}} \beta_{up}' \begin{pmatrix} 1 \\ W_{kl,c,0} \end{pmatrix} g_{li}}_{\text{popularity}} \end{aligned}$$

where, as in Mele (2017), we impose that $\beta_{un} = \beta_{up}$. This specification accounts for gains in direct links, reciprocity, indirect links and "popularity" (individuals derive utility of serving as a "bridge" between agents). The restriction $\beta_{un} = \beta_{up}$ is an identification assumption in Mele's setting, where networks are draws from the model's stationary distribution. It is not required in our setup; though we enforce it for the sake of comparability. The difference in preference shocks is drawn from a logistic distribution.

We estimate our model using a two-step approach. In the first step, we estimate τ_0 through (4). We then use the likelihood-free procedure outlined in Section 4.2 and Appendix F to approximate for the posterior of preference- and matching-related parameters, where we take the first-step estimate of τ_0 as given[32].

Our first-step estimate of τ_0 leads to 76 rounds. The estimator violates the bound $\hat{\tau} \leq N_c(N_c - 1)$ for 80 out of 161 networks in our sample, which is somewhat reassuring[33].

In the second step, we use independent zero-mean normals with a one-ninth standard deviation as priors[34]; and set the number of simulations to $100,000$. We aim for an acceptance rate of 1%; choose the whole data as our vector of statistics; and compute the distance between simulated and observed networks as $\|\{\tilde{G}_c^{T_1}\}_{c=1}^C - \{G_c^{T_1}\}_{c=1}^C\| = \sum_{c=1}^C \|\tilde{G}_c^{T_1} - G_c^{T_1}\|_1$, i.e. we count the number of differing edges in the artificial and observed datasets. We consider a "smooth" rejection rule, where a draw is accepted with probability $\phi(\|\{\tilde{G}_c^{T_1}\}_{c=1}^C - \{G_c^{T_1}\}_{c=1}^C\|/\epsilon)/\phi(0)$. with ϕ being the pdf of a standard normal. We also employ importance sampling, where the "efficient" – in terms of minimising the variance of the approximation to the posterior mean – proposal distribution is chosen according to the algorithm in Li and Fearnhead (2018).

Table 2 summarises our results. We find that most mean estimates of utility parameters associated with covariates are negative, which seems to suggest homophily in preferences is pervasive and not restricted to direct links. Mean estimates of matching parameters are also mostly negative. In particular, we notice that the posterior probability of the coefficient associated with our instrument being negative is over 99.9%, which is reassuring. We also find high posterior probabilities of a negative effect for the role of age in the direct and indirect components of utility; and for the role of cognitive skills in the reciprocal part of utility. Interestingly, the reciprocal component in utility associated with socioemotional

[32]From a frequentist perspective, we can motivate our estimator by appealing to some Bernstein-von-Mises theorem (Vaart, 1998, chapter 10) which ensures posterior consistency as $C \to \infty$. From a Bayesian point of view, our estimator imposes a degenerate prior on τ_0 at the frequentist estimator (4).

[33]Identification of τ_0 in our setting extends immediately to the case where the number of agents in a network, N_c, is assumed to be random, provided that $\mathbb{P}[\tau_0 \leq N_c(N_c - 1)] > 0$.

[34]Our choice of standard deviation is motivated by the magnitude of reduced-form coefficients in Table 1

skills appears to be "heterophilious".

Table 2: Posterior estimates - Optimal proposal method

	Mean	Q 0.025	Q 0.975	Prob < 0
Meeting process				
distance in age	0.0478	-0.0573	0.1409	0.1905
distance in gender	0.0255	-0.1310	0.1526	0.2582
distance in cognitive skills	0.0255	-0.0360	0.0784	0.1503
distance in conscientiousness	-0.0059	-0.0877	0.0756	0.5703
distance in neuroticism	-0.0157	-0.1034	0.0766	0.6642
g_ij	-0.0254	-0.1208	0.0487	0.7685
(1-g_ij)*distance in classlist	-0.0967	-0.1566	-0.0342	0.9993
Utility – Direct Links				
intercept	-0.0327	-0.0972	0.0193	0.8438
distance in age	-0.0908	-0.1569	-0.0120	0.9881
distance in gender	-0.0119	-0.1409	0.0887	0.5988
distance in cognitive skills	-0.0118	-0.1000	0.0765	0.5328
distance in conscientiousness	-0.0256	-0.1313	0.0834	0.6768
distance in neuroticism	-0.0254	-0.1093	0.0589	0.6377
Utility – Mutual Links				
intercept	-0.0423	-0.1185	0.0551	0.8452
distance in age	-0.0405	-0.1675	0.0617	0.6549
distance in gender	-0.0239	-0.0916	0.0515	0.7553
distance in cognitive skills	-0.0651	-0.1341	0.0025	0.9590
distance in conscientiousness	0.0431	-0.0033	0.0839	0.0413
distance in neuroticism	0.0454	-0.0173	0.1081	0.0835
Utility – Indirect links/Popularity				
intercept	-0.0190	-0.0740	0.0310	0.8317
distance in age	-0.0859	-0.1633	-0.0211	0.9943
distance in gender	-0.0145	-0.0695	0.0532	0.6966
distance in cognitive skills	-0.0307	-0.1019	0.0498	0.7346
distance in conscientiousness	-0.0300	-0.1138	0.0468	0.7906
distance in neuroticism	-0.0266	-0.0871	0.0353	0.7876

Next, we proceed to counterfactual exercises. We consider the evolution of networks, starting from their baseline value, under three different sequences of matching parameters: (i) when these are kept at their estimated value (*base* case); (ii) when random unbiased matching is imposed across networks[35] (*random matching* case); (iii) when, keeping grade and classroom sizes in schools fixed, we track students according to their cognitive skills (*tracking* case); and (iv) when, upon meeting, students form a relationship with probability 1/2 (*random friendship* case).

Table 3 reports posterior means and 95% credible intervals of the projection coefficients of edge indicators at the followup period ($g_{ij,c,1}$) on an intercept and our main controls *at the baseline*. Compared with the observed data, magnitudes in the base case are broadly in line with the frequentist reduced form (column (2) in Table 1). Nonetheless, we note our model appears to overstate the role of homophily in age and understate the role of gender. Comparing the first and second columns, one notices imposing random unbiased matching slightly increases observed homophily patterns in cognitive skills. This is somewhat intuitive, as the estimated matching function displays heterophily with respect to this trait. Moving on to the third column, we note that tracking leads to a weak reduced-form estimate of homophily in cognitive skills, which is expected as students now interact in homogeneous groups. It also leads to a weaker pattern in the gender coefficient, which may be due to correlation of this attribute with cognitive skills at the baseline[36]. Finally, we note that the imposing random friendship formation eliminates homophily in age, as expected.

[35] In other words, for individuals i and j in classroom c, we set $\rho((i,j), g, X_c) = \frac{1}{N_c(N_c-1)}$

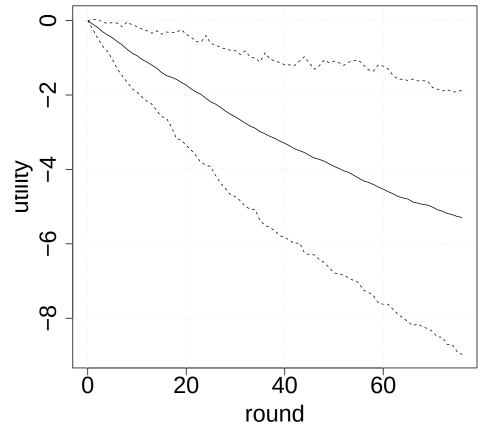
[36] Women have, on average, 0.0185 more points in cognitive skills at the baseline than men, and this difference is statistically significant at the 1% level.

Table 3: Homophily measures under base counterfactual policies

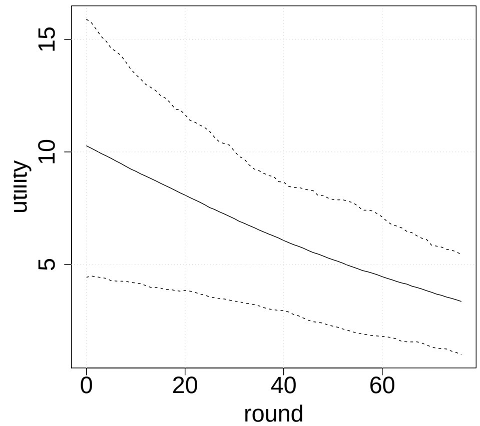
	Base case	Random matching	Tracking	Random friendship
	Regression coefficients			
distance in classlist	-0.0068	-0.0016	-0.0124	-0.0096
	[-0.0097;-0.0035]	[-0.0024;-6e-04]	[-0.0156;-0.0086]	[-0.0122;-0.0054]
distance in age	-0.0146	-0.0161	-0.0152	0.0011
	[-0.0196;-0.0085]	[-0.0219;-0.0103]	[-0.024;-0.0053]	[-0.0047;0.0069]
distance in gender	-0.0921	-0.1077	-0.0421	-0.0938
	[-0.1068;-0.0775]	[-0.1197;-0.0958]	[-0.0532;-0.0309]	[-0.1055;-0.0849]
distance in cognitive skills	-0.0955	-0.1342	0.0322	-0.1096
	[-0.1485;-0.0446]	[-0.1982;-0.0675]	[-0.0468;0.1613]	[-0.1557;-0.0529]
distance in conscientiousness	-0.0122	-0.0117	-0.0079	-0.0112
	[-0.0259;0.0027]	[-0.0267;0.0025]	[-0.0249;0.0053]	[-0.0215;0.0013]
distance in neuroticism	-0.0111	-0.0106	-0.0102	-0.0086
	[-0.0201;-0.0039]	[-0.0215;-0.0023]	[-0.0205;0.005]	[-0.0251;0.0046]
	Degree summary statistics			
Avg. degree	0.1935	0.23	0.1709	0.2479
	[0.1687;0.2241]	[0.207;0.261]	[0.1515;0.1967]	[0.2136;0.2802]
Var. degree	0.1559	0.177	0.1416	0.1862
	[0.1402;0.1739]	[0.1641;0.1929]	[0.1285;0.158]	[0.168;0.2017]

Figure 1 compares the evolution of aggregate utility in the base case with each of our counterfactual scenarios. We plot posterior means and 95% credible intervals of the aggregate utility index, $\sum_{c=1}^{C}\sum_{i=1}^{N_c} u_i(g_c^t, X_c)$, in the counterfactual scenario **minus** the same index in the base case, for each round from the baseline to the followup period. We normalise the difference in indices by the number of students in our sample *times* 0.0908, the absolute value of the direct-link homophily in age mean coefficient (Table 2); so results can be interpreted as the required change in the age distance of direct links each student should receive in the base case so they are indifferent between policies (without taking into account spillovers on the remaining components of utility). One notices that imposing random matching leads to a lower trajectory in aggregate utility over the school year. We also see that tracking leads to an improvement in welfare, though this benefit diminishes as time passes by. Finally, random friendship formation leads to lower welfare.

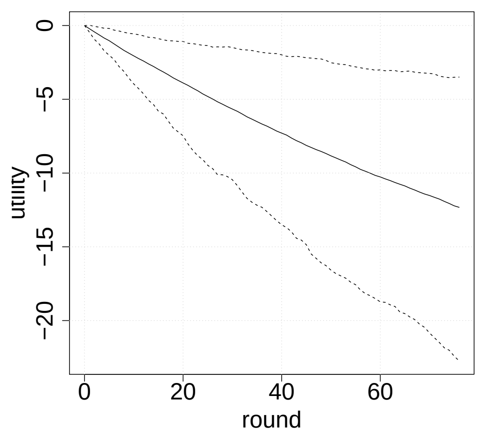
Figure 1: Aggregate utility under counterfactual policies



(a) Random matching



(b) Tracking



(c) Random friendship

6 Concluding remarks

In this article, we studied identification and estimation of a network formation model that distinguishes between homophily that is due to preferences; and homophily that is due to meeting opportunities. The model builds upon the algorithm in Mele (2017) by allowing for rather general classes of utilities and meeting processes. It is also well-grounded on the theoretical literature of network formation (Jackson and Watts, 2002; Jackson, 2010).

We provided identification results in the case a large-support "instrument" is included in preferences (meeting process) and excluded from the meeting process (preferences); and two periods of data from many networks are available. We also discussed a Bayesian estimation procedure that bypasses direct evaluation of the model likelihood, a task which can be computationally unfeasible even for a moderate number of rounds of the network algorithm.

In the applied section of our article, we studied network formation in elementary schools in Northeastern Brazil. Our results suggest that removing biases in the meeting process may actually increase homophily in some dimensions. We also find that tracking students according to their cognitive skills leads to improved welfare, though the benefits diminish over time. Our counterfactual exercise further shows eliminating biases in meeting opportunities produces a lower path of aggregate utility, though this evidence should be taken with caution due to somewhat wide estimated credible intervals.

As emphasised in Section 3.3, analysing how restrictions on the relationship between the network structure and pay-offs may enable point-identification in our setting is an open question which may further enhance the model's applicability – especially if it allows us to dispense with the exclusion restrictions currently required to identify the model. Another interesting topic for future research is the study of our model under single-network asymptotics, which may be more suitable in some applied settings.

References

Aguirregabiria, V. and P. Mira (2010). Dynamic discrete choice structural models: A survey. *Journal of Econometrics 156*(1), 38 – 67. Structural Models of Optimization Behavior in Labor, Aging, and Health.

Amemiya, T. (1985). *Advanced Econometrics.* Harvard University Press.

Badev, A. (2017). Discrete Games in Endogenous Networks: Equilibria and Policy. Working paper.

Bajari, P., H. Hong, and S. P. Ryan (2010). Identification and estimation of a discrete game of complete information. *Econometrica 78*(5), 1529–1568.

Blum, M. G. B. (2010). Approximate bayesian computation: A nonparametric perspective. *Journal of the American Statistical Association 105*(491), 1178–1187.

Bramoullé, Y., S. Currarini, M. O. Jackson, P. Pin, and B. W. Rogers (2012). Homophily and long-run integration in social networks. *Journal of Economic Theory 147*(5), 1754 – 1786.

Chandrasekhar, A. (2016). Econometrics of network formation. In *The Oxford Handbook of the Economics of Networks*, pp. 303–357. Oxford University Press.

Chandrasekhar, A. and M. O. Jackson (2016). A network formation model based on subgraphs.

Christakis, N. A., J. H. Fowler, G. W. Imbens, and K. Kalyanaraman (2010, May). An empirical model for strategic network formation. Working Paper 16039, National Bureau of Economic Research.

Currarini, S., M. O. Jackson, and P. Pin (2009). An economic model of friendship: Homophily, minorities, and segregation. *Econometrica 77*(4), 1003–1045.

Currarini, S., M. O. Jackson, and P. Pin (2010). Identifying the roles of race-based choice and chance in high school friendship network formation. *Proceedings of the National Academy of Sciences 107*(11), 4857–4861.

de Paula, A. (2016). Econometrics of network models. CeMMAP working papers CWP06/16, Centre for Microdata Methods and Practice, Institute for Fiscal Studies.

de Paula, Á́., S. Richards-Shubik, and E. Tamer (2018). Identifying preferences in networks with bounded degree. *Econometrica 86*(1), 263–288.

Frazier, D. T., G. M. Martin, C. P. Robert, and J. Rousseau (2018). Asymptotic properties of approximate bayesian computation. *Biometrika 105*(3), 593–607.

Goldsmith-Pinkham, P. and G. W. Imbens (2013). Social networks and the identification of peer effects. *Journal of Business & Economic Statistics 31*(3), 253–264.

Graham, B. S. (2016, April). Homophily and transitivity in dynamic network formation. Working Paper 22186, National Bureau of Economic Research.

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

Higham, N. J. and L. Lin (2011). On pth roots of stochastic matrices. *Linear Algebra and its Applications 435*(3), 448 – 463. Special Issue: Dedication to Pete Stewart on the occasion of his 70th birthday.

Hobert, J. P. (2011). Chapter 10 the data augmentation algorithm: Theory and methodology. In *Handbook of Markov Chain Monte Carlo*, pp. 253–294. CRC Press.

Horn, R. A. and C. R. Johnson (2012). *Matrix Analysis*. Cambridge University Press.

Hsieh, C.-S. and L. F. Lee (2016). A social interactions model with endogenous friendship formation and selectivity. *Journal of Applied Econometrics 31*(2), 301–319. jae.2426.

Jackson, M. (2010). *Social and Economic Networks*. Princeton University Press.

Jackson, M. O. and A. Watts (2002). The evolution of social and economic networks. *Journal of Economic Theory 106*(2), 265 – 295.

Li, Q. and J. S. Racine (2006). *Nonparametric Econometrics: Theory and Practice*. Princeton University Press.

Li, W. and P. Fearnhead (2018). On the asymptotic efficiency of approximate bayesian computation estimators. *Biometrika 105*(2), 285–299.

McFadden, D. (1973). Conditional logit analysis of qualitative choice behaviour. In P. Zarembka (Ed.), *Frontiers in Econometrics*, pp. 105–142. New York, NY, USA: Academic Press New York.

Mele, A. (2017). A structural model of dense network formation. *Econometrica 85*(3), 825–850.

Newey, W. K. and D. McFadden (1994). Chapter 36 large sample estimation and hypothesis testing. In *Handbook of Econometrics*, Volume 4, pp. 2111 – 2245. Elsevier.

Newey, W. K. and R. J. Smith (2004). Higher order properties of gmm and generalized empirical likelihood estimators. *Econometrica 72*(1), 219–255.

Norris, J. R. (1997). *Markov Chains*. Cambridge Series in Statistical and Probabilistic Mathematics. Cambridge University Press.

Pinto, C. C. d. X. and V. P. Ponczek (2017). The Building Blocks of Skill Development. Unpublished.

Rothenberg, T. J. (1971). Identification in parametric models. *Econometrica 39*(3), 577–591.

Sheng, S. (2014). A structural econometric analysis of network formation games. `http://www.econ.ucla.edu/people/papers/Sheng/Sheng626.pdf`.

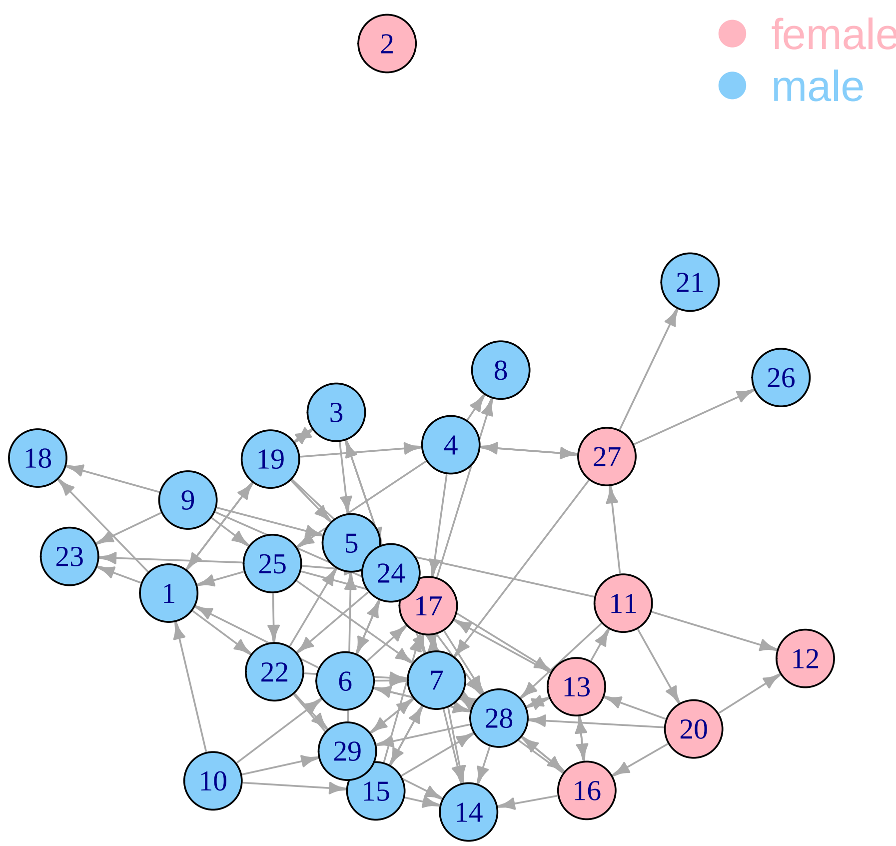
Sisson, S. A. and Y. Fan (2011). Chapter 12 likelihood-free mcmc. In *Handbook of Markov Chain Monte Carlo*, pp. 313–338. CRC Press.

Tamer, E. (2003). Incomplete simultaneous discrete response model with multiple equilibria. *The Review of Economic Studies 70*(1), 147–165.

Vaart, A. W. v. d. (1998). *Asymptotic Statistics*. Cambridge Series in Statistical and Probabilistic Mathematics. Cambridge University Press.

A Application: auxiliary tables

Figure 2: Example: a 3rd grade baseline classroom



Note: The figure presents a 3rd grade classroom network from our baseline data. Numbered circles represent students. An arrow stemming from circle "x" to "y" denotes student "x" nominated "y" as a friend in early 2014.

Table 4: Pairwise distance in covariates – Summary statistics

Statistic	N	Mean	St. Dev.	Min	Pctl(25)	Pctl(75)	Max
edge (baseline)	17,736	0.172	0.377	0	0	0	1
edge (followup)	17,736	0.179	0.383	0	0	0	1
distance in classlist (baseline)	17,736	8.009	5.553	1	3	12	30
distance in age (years) (baseline)	17,736	0.827	0.908	0.000	0.249	1.041	6.633
distance in gender (baseline)	17,736	0.502	0.500	0	0	1	1
distance in cognitive skills (baseline)	17,736	0.099	0.081	0.000	0.035	0.144	0.530
distance in conscientiousness (baseline)	17,736	0.568	0.485	0.000	0.207	0.798	3.275
distance in neuroticism (baseline)	17,736	0.676	0.538	0.000	0.249	0.976	3.491

B Identification with one period of data from the stationary distribution

In this Appendix, we explore identification in a context where we have access to a sample of C networks stemming from the stationary distribution of the game described in Section 2. In particular, we have access to a sample $\{X_c, G_c\}_{c=1}^{C}$. Note from Remark 2.2 that one may interpret the stationary distribution as a long-run distribution. If one assumes observed network data (G_c) was drawn from the (conditional) stationary distribution, then the (conditional) stationary distribution $\pi(X)$ is identified. Notice that, in this context, identification of the transition matrix $\Pi(X)$ is a *necessary* condition for identification of $((u_i)_{i=1}^{N}, \rho)$, the objects of interest. We thus propose to analyse identification of $\Pi(X)$. In particular, we explore identification of $\Pi(X)$ without imposing further restrictions. In light of this, and without loss of generality, we may essentially view X as nonstochastic throughout the remainder of this section and suppress dependence of $\Pi(X)$ on X by writing Π,

In our setting, the identification problem (of Π) reduces to providing conditions under which no other transition matrix $\tilde{\Pi} \in \mathcal{S}$ is observationally equivalent to Π; where $\mathcal{S}$ is the admissible (by the model) set of Markov chains. In other words, Π is identified if:

$$\forall\ \tilde{\Pi} \in \mathcal{S} \quad (I - \tilde{\Pi}')\pi = 0 \implies \tilde{\Pi} = \Pi \tag{6}$$

If $\mathcal{S}$ were the set of all row-stochastic matrices, Π would clearly not be identified, as $\mathbb{I}_{2^{N(N-1)} \times 2^{N(N-1)}}$ is observationally equivalent. But $\mathcal{S}$ is not the set of all Markov matrices. Indeed, the model imposes restrictions on the set of admissible Markov matrices. These are summarised by (2). As we do not impose further restrictions on utilities and the matching function, $\mathcal{S}$ *is the set of all* $2^{N(N-1)} \times 2^{N(N-1)}$ *row-stochastic matrices with strictly positive entries* Π_{gw} *for all* $g \in \mathcal{G}$, $w \in N(g) \cup \{g\}$, *and 0 otherwise.*

Do the restrictions implied by the model identify Π? The next lemma is a negative result.

Lemma B.1 (Non-identification)**.** *Under Assumption 2.1 and Assumption 2.2, and if* $F_\epsilon(e_0, e_1) = \exp[-\exp(e_0) - \exp(e_1)]$ *(i.e.* $(\epsilon(0), \epsilon(1))$ *are independent EV type 1), then the model is not identified.*

Proof. Fix $\Pi_0 \in \mathcal{S}$ and let π_0 be the (unique) solution to $(I - \Pi_0')\pi_0 = 0$, $\pi_0'\iota = 1$. The proof consists in presenting a family of observationally equivalent versions of the model in Mele (2017), albeit in a more general setting. Consider a family of utilities $(u_i)_{i=1}^N$ where $u_i(g, X) = \ln(\pi_0(g))$ for all $i \in \mathcal{I}$ and all $g \in \mathcal{G}$. These utilities are well defined, as $\pi_0 >> 0$ from Remark 2.3. Moreover, fix some arbitrary ρ satisfying (i) Assumption 2.1; and (ii) $\rho((i,j), [0, g_{-ij}], X) = \rho((i,j), [1, g_{-ij}], X)$ for all $(i,j) \in \mathcal{M}$, $g \in \mathcal{G}$ (the matching probability does not depend on the existence of a link between i and j). These choices satisfy Assumption 2.1 and Assumption 2.2. Therefore, there exists a unique stationary distribution $\tilde{\pi}$ associated with the chain $\tilde{\Pi}$ from this game. Further notice that the family of utilities admits a *potential function* $Q : \mathcal{G} \times \mathcal{X} \mapsto \mathbb{R}$ satisfying $Q([1, g_{-ij}], X) - Q([0, g_{-ij}, X]) = u_i([1, g_{-ij}], X) - u_i([0, g_{-ij}], X)$ for all $(i,j) \in \mathcal{M}$, $g \in \mathcal{G}$. Indeed, $Q(g, X) = \ln(\pi_0(g))$ is a potential function for this class of utilities. But then, this distribution has a closed form expression: $\tilde{\pi}(g) = \frac{\exp(Q(g,X))}{\sum_{w \in \mathcal{G}} \exp(Q(w,X))} = \pi_0(g)$. To see this, one can follow the argument in Theorem 1 of Mele (2017) and verify that this distribution satisfies the flow balancedness condition $\tilde{\pi}_g \tilde{\Pi}_{gw} = \tilde{\pi}_w \tilde{\Pi}_{wg}$ for all $g, w \in \mathcal{G}$[37]. But then the

[37]See Appendix C.

stationary distribution does not depend on the matching probabilities, so the Markov matrix is not identified, as it is always possible to choose ρ satisfying $(i)-(ii)$ s.t. $\tilde{\Pi} \neq \Pi_0$ (and $\pi_0 = \tilde{\pi}$ follows, as we saw). $\square$

We interpret this result as suggestive of the need to impose further restrictions in order to identify Π. Note that verification of the flow balancedness condition relies crucially on the functional form of the distribution function for the error term. One approach would then be to restrict analysis to different distributions. We do not follow this approach, as it is not grounded in knowledge regarding the social interactions being analysed (identification by functional form). Moreover, it is immediate to see the identification at infinity strategy discussed in Section 3.2 does not bring additional identifying power in the setting of Lemma B.1 without further restrictions on the matching function and/or utilities. In particular, it would require restricting $((u_i)_{i=1}^N, \rho)$ to classes where the assumptions in Mele (2017) regarding the matching function and/or utilities do not hold. Since this would restrict the generality and applicability of the model, we refrain from further analysing identification with one period of data from the stationary distribution.

C Flow balancedness condition

We also show the flow balancedness condition in the model of Mele (2017) holds:

Claim C.1 (Mele (2017))**.** *In the model of Mele (2017), the flow-balancedness condition* $\tilde{\pi}_g \tilde{\Pi}_{gw} = \tilde{\pi}_w \tilde{\Pi}_{wg}$ *for all* $g, w \in \mathcal{G}$ *holds.*

Proof. Fix $g, w \in \mathcal{G}$. Note that if g and w differ in more than one node, the condition trivially holds. This is also the case when $g = w$. Consider then the case where g and w

differ exactly in one node, say $g_{ij} = 1$ and $w_{ij} = 0$ for some $i, j \in \mathcal{I}$. We then have that

$$\begin{aligned}
\tilde{\pi}_g \tilde{\Pi}_{gw} &= \frac{\exp(Q(g,X))}{\sum_{s\in\mathcal{G}} \exp(Q(s,X))} \times \rho((i,j),g,X) \frac{\exp(Q(w,X)-Q(g,X))}{1+\exp(Q(w,X)-Q(g,X))} = \\
&= \frac{\exp(Q(g,X)-Q(w,X)+Q(w,X))}{\sum_{s\in\mathcal{G}} \exp(Q(s,X))} \times \rho((i,j),g,X) \frac{\exp(Q(w,X)-Q(g,X))}{1+\exp(Q(w,X)-Q(g,X))} \\
&= \frac{\exp(Q(w,X))}{\sum_{s\in\mathcal{G}} \exp(Q(s,X))} \times \rho((i,j),g,X) \frac{1}{1+\exp(Q(w,X)-Q(g,X))} = \\
&= \frac{\exp(Q(w,X))}{\sum_{s\in\mathcal{G}} \exp(Q(s,X))} \times \rho((i,j),g,X) \frac{\exp(Q(g,X)-Q(w,X))}{1+\exp(Q(g,X)-Q(w,X))} \overset{\text{(ii)}}{=} \\
&\overset{\text{(ii)}}{=} \frac{\exp(Q(w,X))}{\sum_{s\in\mathcal{G}} \exp(Q(s,X))} \times \rho((i,j),w,X) \frac{\exp(Q(g,X)-Q(w,X))}{1+\exp(Q(g,X)-Q(w,X))} = \tilde{\pi}_w \tilde{\Pi}_{gw}
\end{aligned} \tag{7}$$

□

D Proof of Claim 3.1

Observe that $\Pi(\gamma)$ is written as:

$$\Pi(\gamma) = \begin{array}{c} \begin{array}{cccc} (0,0) & (1,0) & (0,1) & (1,1) \end{array} \\ \begin{pmatrix} \rho_{12}(0,0)F_{12}((1,0),(0,0)) + \rho_{21}(0,0)F_{21}((0,1),(0,0)) & \rho_{12}(0,0)F_{12}((0,0),(1,0)) & \rho_{21}(0,0)F_{21}((0,0),(0,1)) & 0 \\ \rho_{12}(1,0)F_{12}((1,0),(0,0)) & \rho_{12}(1,0)F_{12}((0,0),(1,0)) + \rho_{21}(1,0)F_{21}((1,1),(1,0)) & 0 & \rho_{21}(1,0)F_{21}((1,0),(1,1)) \\ \rho_{21}(0,1)F_{21}((0,1),(0,0)) & 0 & \rho_{12}(0,1)F_{12}((1,1),(0,1)) + \rho_{21}(1,1)F_{21}((0,0),(0,1)) & \rho_{12}(0,1)F_{12}((0,1),(1,1)) \\ 0 & \rho_{21}(1,1)F_{21}((1,1),(1,0)) & \rho_{12}(1,1)F_{12}((1,1),(0,1)) & \rho_{12}(1,1)F_{12}((0,1),(1,1)) + \rho_{21}(1,1)F_{21}((1,0),(1,1)) \end{pmatrix} \end{array} \begin{array}{c} \\ (0,0) \\ (1,0) \\ (0,1) \\ (1,1) \end{array}$$

Now suppose there exists $\gamma \neq \tilde{\gamma}$, $\Pi(\gamma) = \Pi(\tilde{\gamma})$. Suppose $\rho_{21}(0,0) > \widetilde{\rho_{21}(0,0)}$. The argument is symmetric for the remaining parameters. Since objects are observationally equivalent, it must be that $F_{21}((0,0),(0,1)) < \widetilde{F_{21}((0,0),(0,1))}$, which in its turn yields $F_{21}((0,1),(0,0)) > \widetilde{F_{21}((0,1),(0,0))}$. This implies $\rho_{21}(0,1) < \widetilde{\rho_{21}(0,1)}$; consequently, $\rho_{12}(0,1) > \widetilde{\rho_{12}(1,0)}$ and $F_{12}((0,1),(1,1)) < \widetilde{F_{12}((0,1),(1,1))}$ thereafter. But then $F_{12}((1,1),(0,1)) < \widetilde{F_{12}((1,1),(0,1))}$, leading to $\rho_{12}(1,1) > \widetilde{\rho_{12}(1,1)}$, $\rho_{21}(1,1) < \widetilde{\rho_{21}(1,1)}$, $F_{21}((1,1),(1,0)) > \widetilde{F_{21}((1,1),(1,0))}$. Proceeding analogously, we will get $\rho_{21}(1,0) > \widetilde{\rho_{21}(0,1)}$, $F_{12}((0,0),(1,0)) < \widetilde{F_{12}((0,0),(1,0))}$ and finally $\rho_{12}((0,0)) > \widetilde{\rho_{12}(0,0)}$. But we also had $\rho_{21}(0,0) > \widetilde{\rho_{21}(0,0)}$, leading to $1 > 1$, a contradiction.

E Identification with exclusion restriction on matching function

In this section, we analyse how our main identification result (Proposition 3.1) would change if large support variables were included in the matching function (ρ), but not in utilities. Fix $g \in \mathcal{G}$, $w \in N(g)$, $g_{ij} \neq w_{ij}$. We start by establishing the following claim:

Claim E.1. *Under a limit which drives $\rho_{ij}(g) \to 1$ and $\rho_{ij}(w) \to 1$; but leaves $F_{ij}(g,w)$ unaltered, we have:*

$$\lim_{t^*} (\Pi^{\tau_0})_{gw} = F_{ij}(g,w)$$

Proof. The case where $\tau_0 = 1$ is immediate, since $\lim_{t^*}(\Pi^{\tau_0})_{gw} = F_{ij}(g,w)$ follows directly from (2). Suppose $\tau_0 > 1$. Recall that:

$$(\Pi^{\tau_0})_{gw} = \sum_{m \in N(w) \cup \{w\}} (\Pi^{\tau_0 - 1})_{gm} \Pi_{mw}$$

Using a similar argument as in Proposition 3.1, we can show the limit in the statement of the Claim is such that $\lim_{t^*}(\Pi^{\tau_0-1})_{gm} \to 0$ for all $m \in N(w) \setminus \{g\}$[38]. We are thus left with:

$$\lim_{t^*}(\Pi^{\tau_0})_{gw} = \lim_{t^*}(\Pi^{\tau_0-1})_{gw}\Pi_{ww} + \lim_{t^*}(\Pi^{\tau_0-1})_{gg}\Pi_{gw} = (\lim_{t^*}(\Pi^{\tau_0-1})_{gw} + \lim_{t^*}(\Pi^{\tau_0-1})_{gg})F_{ij}(g,w)$$

Next, we note that, under the limit in the statement of the claim:

$$\lim_{t^*}(\Pi^{\tau_0})_{gg} = (\lim_{t^*}(\Pi^{\tau_0-1})_{gw} + \lim_{t^*}(\Pi^{\tau_0-1})_{gg})F_{ij}(w,g)$$

which follows from $(\Pi^{\tau_0})_{gg} = \sum_{m \in N(g) \cup \{g\}} (\Pi^{\tau_0-1})_{gm}\Pi_{mg}$ and an argument similar to the previous one. We then get:

[38] If $m \in N(w) \setminus \{g\}$, then $m \in N^2(g)$. Fix a summand in $(\Pi^{\tau_0-1})_{gm}$. If a transition from g to m occurs at pair $(k,l) \neq (i,j)$, the limit in the statement of the Claim drives the summand to 0, If all transitions occur at pair (i,j), a transition from w to some $z \in N(w) \setminus \{g\}$ must occur, for, if not, then either $m = g$ or $m = w$, which is not true. The limit in the statement of the Claim thus drives the summand to zero.

$$\lim_{t^*}(\Pi^{\tau_0})_{gw} + \lim_{t^*}(\Pi^{\tau_0})_{gg} = \lim_{t^*}(\Pi^{\tau_0-1})_{gw} + \lim_{t^*}(\Pi^{\tau_0-1})_{gg}$$

Induction on $\lim_{t^*}(\Pi)_{gw} + \lim_{t^*}(\Pi)_{gg} = 1$ then yields the desired result. $\square$

The previous argument suggests that, if large support variables included in $\rho_{ij}(g)$ and $\rho_{ij}(w)$ – but excluded from $F_{ij}(g,w)$ – may be shifted in a direction that (simultaneously) drives the probability of selecting pair (i,j) under networks w and g to 1; then marginal utilities are identified.

Identification of the matching process in this setting is more intricate, as we require the "feasibility" of a different limit. We consider the case where $\tau_0 \leq N(N-1)$ (Assumption 3.2 holds). Fix some $s \in N^{\tau_0}(g)$ such that $g_{ij} \neq s_{ij}$. Denote by $\mathcal{D} \subseteq \mathcal{M}$ be the set of pairs where g and s differ. We then have:

$$\begin{aligned}
&(\Pi^{\tau_0})_{gs} = \\
&\sum_{(a_1,a_2\ldots a_{\tau_0})\in P(D)} \rho_{a_1}(g)F_{a_1}(g,[1-g_{a_1},g_{-a_1}])\times \\
&\rho_{a_2}([1-g_{a_1},g_{-a_1}])F_{a_2}([1-g_{a_1},g_{-a_1}],[1-g_{a_1},1-g_{a_2},g_{-a_1,-a_2}])\times\ldots \\
&\times\rho_{a_{\tau_0}}([1-g_{a_1},1-g_{a_2}\ldots 1-g_{a_{\tau_0-1}},g_{-a_1,-a_2\ldots-a_{\tau_0-1}}])\times \\
&F_{a_{\tau_0}}([1-g_{a_1},1-g_{a_2}\ldots 1-g_{a_{\tau_0-1}},g_{-a_1,-a_2\ldots-a_{\tau_0-1}}],s)
\end{aligned}$$

where $P(\mathcal{D})$ is the set of all vectors constructed from permutations of the elements in $\mathcal{D}$. If there exists a limit that vanishes all summands not starting on $\rho_{ij}(g)$ (but leaves the latter unchanged); and if it is further feasible to simultaneously drive $\rho_{ij}(g)$ to 1, then a ratio of limits identifies matching probabilities.

F Approximate Bayesian Computation (ABC) algorithm

In this section, we show our likelihood-free algorithm provides an approximation of moments of the posterior distribution. As in our main text, we observe a sample $\{G_c^{T_0}, G_c^{T_1}, X_c\}_{c=1}^{C}$ from the model. As our focus lies on the posterior distribution, we essentially view this sample as fixed (nonstochastic) throughout the remainder of this section. The model parameters are $(\beta,\tau) \in \mathbb{B}\times\mathbb{N}$. The prior density is p_0, and the model likelihood is $\mathbb{P}(\cdot|\mathbf{X}_C;\beta,\tau)$, where $\mathbf{X}_C := \{X_c, G_c^{T_0}\}_{c=1}^C$. Approximate Bayesian Computation requires we be able to draw samples from $\mathbb{P}(\cdot|\mathbf{X}_C;\beta,\tau)$ and compare it with the data $\mathbf{Y}_C := \{G_c^{T_1}\}_{c=1}^C$. In particular, we consider computing (approximations of) moments of the posterior distribution $\mathbb{P}(\cdot|\mathbf{X}_C,\mathbf{Y}_C)$ according to Algorithm 2, where $K(\tilde{\mathbf{Y}}_{Cs},\mathbf{Y}_C;\epsilon)$ is a rescaled kernel and q_0 is a proposal density.

Algorithm 2 Approximating posterior moments

define some tolerance $\epsilon > 0$

define a function $h : \mathbb{B} \times \mathbb{N} \mapsto \mathbb{R}^m$

for $s \in \{1, 2 \ldots S\}$ **do**

draw $(\beta_s, \tau_s) \sim q_0$

generate an artificial sample $\tilde{\mathbf{Y}}_{Cs} \sim \mathbb{P}(\cdot|\mathbf{X}_C;\beta_s,\tau_s)$

accept (β_s,τ_s) with probability $K(\tilde{\mathbf{Y}}_{Cs},\mathbf{Y}_C;\epsilon)$

end for

compute the approximation to the posterior mean of h using the accepted draws according to $\hat{h} := \sum_{s:\text{accepted}} h(\beta_s,\tau_s)w_s / \sum_{s:\text{accepted}} w_s$, where $w_s := p_0(\beta_s,\tau_s)/q_0(\beta_s,\tau_s)$

The next proposition shows that, if we let $\epsilon \to 0$ as $S \to \infty$, our approximation will be consistent for the posterior mean $\mathbb{E}_{\mathbb{P}(\cdot|\mathbf{X}_C,\mathbf{Y}_C)}[h(\cdot)]$.

Proposition F.1. *Suppose that: (1) the prior distribution admits a density p_0 with respect to some measure μ on $\mathbb{B}\times\mathbb{N}$; (2) the proposal density q_0 is such that supp $p_0 \subseteq$ supp q_0; and (3) the map $K : \mathcal{G}^C \times \mathcal{G}^C \times \mathbb{R}_+ \mapsto [0,1]$ is such that (3.i) $K(\mathbf{Y}_1,\mathbf{Y}_2,\cdot)$ is continuous at 0 for all $(\mathbf{Y}_1,\mathbf{Y}_2) \in \mathcal{G}^C \times \mathcal{G}^C$; and (3.ii) $K(\mathbf{Y}_1,\mathbf{Y}_2,0) = \mathbb{1}\{\mathbf{Y}_1 = \mathbf{Y}_2\}$. Then, for any $h : \mathbb{B}\times\mathbb{N} \mapsto \mathbb{R}^m$ such that $\int \|h(\beta,\tau)\| p_0(\beta,\tau) d\mu < \infty$, the approximation $\hat{h}$ in Algorithm 2 is such that $\hat{h} \xrightarrow{p} \mathbb{E}_{\mathbb{P}(\cdot|\mathbf{X}_C,\mathbf{Y}_C)}[h(\cdot)]$ as $\epsilon \to 0$ and $S \to \infty$.*

Proof. Observe that $\hat{h}$ may be written as:

$$\hat{h} = \frac{S^{-1}\sum_{s=1}^S h(\beta_s,\tau_s) w_s \mathbb{1}\{u_s \leq K(\tilde{\mathbf{Y}}_{Js},\mathbf{Y}_C;\epsilon)\}}{S^{-1}\sum_{s=1}^S w_s \mathbb{1}\{u_s \leq K(\tilde{\mathbf{Y}}_{Js},\mathbf{Y}_C;\epsilon)\}} \tag{8}$$

where $\{u_s\}_{s=1}^S$ are iid draws from a Uniform distribution; independently from $\{\beta_s,\tau_s,\tilde{\mathbf{Y}}_{Cs}\}_{s=1}^S$.

We next note the random map $a(\epsilon) = h(\beta_s,\tau_s)w_s\mathbb{1}\{u_s \leq K(\tilde{\mathbf{Y}}_{Cs},\mathbf{Y}_C;\epsilon)\}$ is almost surely continuous at 0, for $\mathbb{P}[\{a \text{ is discontinuous at } 0\}] \leq \mathbb{P}[u_s = 1] = 0$. Moreover, we have that $\mathbb{E}[\sup_{\epsilon\geq 0}\|a(\epsilon)\|] \leq \mathbb{E}[\|h(\beta_s,\tau_s)\|w_s] = \int \|h(\beta,\tau)\|p_0(\beta,\tau)d\mu < \infty$. Then, by applying Lemma 4.3 in Newey and McFadden (1994), we have that, as $S\to\infty$ and $\epsilon\to 0$:

$$S^{-1}\sum_{s=1}^{S} h(\beta_s,\tau_s)w_s\mathbb{1}\{u_s \leq K(\tilde{\mathbf{Y}}_{Js},\mathbf{Y}_C;\epsilon)\} \xrightarrow{p} \mathbb{E}[h(\beta_s,\tau_s)w_s\mathbb{1}\{u_s \leq K(\tilde{\mathbf{Y}}_{Js},\mathbf{Y}_C;0)\}] \quad (9)$$

But we further have that:

$$\begin{aligned}\mathbb{E}[h(\beta_s,\tau_s)w_s\mathbb{1}\{u_s \leq K(\tilde{\mathbf{Y}}_{Js},\mathbf{Y}_C;0)\}] = \mathbb{E}[h(\beta_s,\tau_s)w_s\mathbb{1}\{\tilde{\mathbf{Y}}_{Js}=\mathbf{Y}_C\}] =\\ = \mathbb{E}[h(\beta_s,\tau_s)w_s\mathbb{E}[\mathbb{1}\{\tilde{\mathbf{Y}}_{Js}=\mathbf{Y}_C\}|\beta_s,\tau_s]] = \mathbb{E}[h(\beta_s,\tau_s)w_s\mathbb{P}[\mathbf{Y}_C|\mathbf{X}_C;\beta_s,\tau_s]] =\\ = \int h(\beta_s,\tau_s)\mathbb{P}[\mathbf{Y}_C|\mathbf{X}_C;\beta_s,\tau_s]p_0(\beta_s,\tau_s)d\mu\end{aligned}$$

An analogous argument establishes the denominator converges in probability to $\int \mathbb{P}[\mathbf{Y}_C|\mathbf{X}_C;\beta_s,\tau_s]p_0(\beta_s,\tau_s)d\mu$, which establishes the desired result. □

Examples of maps that satisfy our required property are:

$$K(\mathbf{Y}_1,\mathbf{Y}_2;\epsilon) = \mathbb{1}\{\|\mathbf{Y}_1-\mathbf{Y}_2\|\leq\epsilon\}$$

which corresponds to the "sharp" rejection rule in Algorithm 1 of the main text. A "smooth" alternative is:

$$K(\mathbf{Y}_1,\mathbf{Y}_2;\epsilon) = \begin{cases}\phi(\|\mathbf{Y}_1-\mathbf{Y}_2\|/\epsilon)/\phi(0) & \epsilon>0\\ \mathbb{1}\{\mathbf{Y}_1=\mathbf{Y}_2\} & \epsilon=0\end{cases}$$

where ϕ is the standard normal pdf. This rule is used in our our application (Section 5).