

# Segregation in Social Networks: A Structural Approach\*

Angelo Mele<sup>†</sup>

October 20, 2017

## Abstract

This paper studies racial segregation in schools using data from Add Health. I estimate a structural equilibrium model of friendship formation among students. Preferences depend on direct connections, but also indirect friendships and popularity. I find that students tend to interact with similar people. Homophily goes beyond direct links: students also prefer a racially homogeneous set of indirect friends. I simulate several counterfactual desegregation busing programs, showing that policies that transport minorities to other schools have nonlinear effects on within-school segregation and welfare. In some instances, these interventions increase segregation within schools.

**JEL Codes:** C15, J15, C31

*Keywords:* Racial segregation, Social Networks, Desegregation programs, Bayesian estimation

---

\*Contact: [angelo.mele@jhu.edu](mailto:angelo.mele@jhu.edu). This paper contains a revised version of the empirical implementation of "A structural model of segregation in social networks". I am grateful to Roger Koenker, Ron Laschever, Dan Bernhardt, George Deltas, Matt Jackson, Alberto Bisin, Ethan Cole, Aureo de Paula, Steven Durlauf, Andrea Galeotti, Shweta Gaonkar, Sanjeev Goyal, Dan Karney, Junfu Zhang, Darren Lubotsky, Antonio Mele, Luca Merlino, Tom Parker, Dennis O'Dea, Micah Pollak, Sergey Popov, Sudipta Sarangi, Giorgio Topa, Antonella Tutino and seminar participants at many institutions for helpful comments and suggestions. All remaining errors are mine.

<sup>†</sup>This research uses data from Add Health, a program project designed by J. Richard Udry, Peter S. Bearman, and Kathleen Mullan Harris, and funded by a grant P01-HD31921 from the Eunice Kennedy Shriver National Institute of Child Health and Human Development, with cooperative funding from 17 other agencies. Special acknowledgment is due Ronald R. Rindfuss and Barbara Entwisle for assistance in the original design. Persons interested in obtaining Data Files from Add Health should contact Add Health, The University of North Carolina at Chapel Hill, Carolina Population Center, 123 W. Franklin Street, Chapel Hill, NC 27516-2524 ([addhealth@unc.edu](mailto:addhealth@unc.edu)). No direct support was received from grant P01-HD31921 for this analysis.

# 1 Introduction

Social networks are important determinants of individuals' socioeconomic performance. An increasing amount of evidence shows that the number and composition of social ties affects employment prospects, school performance, risky behavior, adoption of new technologies, diffusion of information and health outcomes.<sup>1</sup> The structure of social ties is endogenous: individuals choose their peers and friends according to their socioeconomic characteristics and their relationships. As a consequence, in socially generated networks the agents are likely to interact with similar individuals (homophily), segregating along socioeconomic attributes.<sup>2</sup>

This paper estimates a structural model of equilibrium network formation with heterogeneous agents, to understand the determinants of racial segregation in social networks. Preferences are defined over individual socioeconomic characteristics and network structures. In each period a random player is selected from the population and he meets another agent, according to a meeting technology. Upon meeting, the player has the opportunity to revise his linking strategy: a link is created (or maintained, if already in place) if and only if the utility of the additional link is positive.

Mele (2017a) proves and characterizes the existence of a *unique stationary equilibrium* for this model, which provides the likelihood of observing a specific network architecture in the long run. I estimate the posterior distribution of the structural parameters, using a Bayesian approach. The main challenge is that the model's likelihood is proportional to an intractable normalizing constant, that cannot be evaluated or approximated with precision. To overcome this problem, I use a Markov Chain Monte Carlo algorithm that generates samples from the posterior distribution without evaluating the likelihood.

Using this theoretical framework, I study segregation in school friendship networks, using data from the *National Longitudinal Study of Adolescent Health* (Add Health). This unique database contains detailed information on friendship networks of students enrolled in a representative sample of US schools. The final sample includes 14 high schools with a total of 1139 students.<sup>3</sup> I find that race, gender and grade are important determinants of network formation in schools. There is overwhelming evidence of homophily: students tend to interact and form social ties with similar people, other things being equal. These estimates control for the structure of the network. Furthermore, I find that homophily effects extend well beyond direct links: for example, students also prefer an homogeneous racial composition of friends of friends.

This model provides useful guidance to policymakers who care about promoting policies that affect the structure of the network. I use the estimated model to predict how a change

---

<sup>1</sup>For example, see the contributions of Topa (2001); Laschever (2009); Cooley (2010); De Giorgi et al. (2010); Nakajima (2007); Bandiera and Rasul (2006); Conley and Udry (2010); Golub and Jackson (2011); Acemoglu et al. (2011).

<sup>2</sup>See Currarini et al. (2009, 2010), De Marti and Zenou (2009), Echenique et al. (2006).

<sup>3</sup>I use only the schools from the *saturated sample*. The sampling scheme of Add Health involved in-school interviews for all the students. A subsample of 20745 students was also interviewed at home, to collect detailed individual information. The saturated sample contains schools for which both interviews were administered to *each* student enrolled. Therefore this sample does not contain any missing information about individual controls. This is not the case for most schools in Add Health.

in the composition of the student population affects the structure of the network and welfare.<sup>4</sup> As an example, I consider a desegregation program that affects two schools in the sample, one with an enrolled population of 98% Whites and the other with 96% African Americans. The simulations consist of alternative re-assignments of students across schools; the outcomes of interest in these counterfactuals are the average segregation in the friendship network within-schools, and the welfare in the new stationary equilibrium of the model.

I find that a desegregation program that equalizes the racial composition across schools is not welfare maximizing and may promote higher segregation within schools. Furthermore, a desegregation plan may have different effects on each school, increasing segregation in one while decreasing it in the other. I also compare the results of these policy experiments with results generated using a model with a simpler preference structure, where individuals care only about their direct links. The predictions of such model are quite different from the full structural model.

The estimation of the posterior distribution is complicated because the likelihood of this model is proportional to a normalizing constant that cannot be evaluated or approximated with precision. Indeed, a state-of-the-art supercomputer would take several years to evaluate the likelihood at a single parameter value. This feature prevents the use of traditional Markov Chain Monte Carlo schemes, e.g. Metropolis-Hastings, that evaluate the likelihood at each iteration. To circumvent this problem, To circumvent this issue, I use a Markov Chain Monte Carlo algorithm that samples from the posterior distribution of the parameters without evaluating the likelihood. The algorithm belongs to the class of *exchange algorithms*, first proposed by [Murray et al. \(2006\)](#) to generate samples from posterior distributions with similar intractable likelihoods.<sup>5</sup>

The idea behind the sampler is as follows. At each iteration we perform a double Metropolis-Hastings step. First, we propose a new parameter vector as in the standard Metropolis-Hastings scheme; second, we draw a new network from the stationary equilibrium of the model at the proposed parameter, using a second Metropolis-Hastings algorithm. If the network generated by the latter simulation is similar to the network observed in the data, then the parameter is accepted with high probability; viceversa, if the simulated network is very different from the data, the probability of acceptance is lower. The similarity between networks is measured using likelihood ratios at the current and proposed parameters.<sup>6</sup> The intuition is that if the observed and simulated networks are similar, the proposed parameter is very likely to generate the data as a draw from the stationary equilibrium. The details about convergence and implementation are shown in Appendix B of [Mele \(2017a\)](#). In the empirical application I use an extension of the algorithm which makes use of parallel computing techniques, allowing estimation of the model using multiple school network data.

This paper contributes to two strands of literature. The empirical network literature

---

<sup>4</sup>Alternatively, the model could be used as a guide for the design of randomized experiments that modify students assignments.

<sup>5</sup>See also [Liang \(2010\)](#), [Caimo and Friel \(2011\)](#), [Mele \(2017a\)](#), [Liang et al. \(2010\)](#) for additional details.

<sup>6</sup>While the evaluation of the likelihood is infeasible, it is always possible to evaluate likelihood ratios computed at the same parameter value, at different network configurations.

has developed models and methods of identification and estimation for strategic network formation models ([Jackson \(2008\)](#), [DePaula \(forthcoming\)](#), [Graham \(2017\)](#), [Chandrasekhar \(2016\)](#)). I contribute to this literature, by estimating the model developed in [Mele \(2017a\)](#) using detailed network data from US high schools, showing that racial homophily is pervasive.<sup>7</sup> This result is consistent with recent evidence on segregation ([Currarini et al. \(2009\)](#), [Boucher \(2015\)](#), [Boucher and Mourifie \(forthcoming\)](#), [Mele \(2017b\)](#)). I also find racial homophily for indirect connections. Standard structural models of network formation that exclude link externalities in the payoffs are not able to capture this feature (see for example [Graham \(2017\)](#), [Dzemeski \(2017\)](#)), attributing all the homophily to direct links.

The rich dataset used in my estimation partially solves the identification issues highlighted in [Mele \(2017a\)](#). Indeed, the identification of the structural parameters is problematic only when the researcher can observe one single network. In this work, I estimate the model using multiple independent network observations: identification is guaranteed by the theory of exponential families distribution ([Lehman \(1983\)](#)).

The paper also contributes to the literature on racial segregation in schools. This body of work focused on the effects of residential and school segregation on minority socioeconomic outcomes ([Cutler and Glaeser \(1997\)](#), [Echenique and Fryer \(2007\)](#), [Ferrara and Mele \(2011\)](#), [Ananat \(2011\)](#)). Other work has considered the effect of school segregation on educational attainment ([Angrist and Lang \(2004\)](#)). Most studies analyze segregation among schools in a district, but few authors have used more detailed data at the school level to understand the patterns of racial segregation within schools ([Echenique and Fryer \(2007\)](#), [Echenique et al. \(2006\)](#), [Mele \(2017b\)](#), [Boucher \(2015\)](#), [Badev \(2013\)](#)). My approach in this paper provides a structural interpretation of the segregation levels within schools.

My model allows me to simulate counterfactual policies that could improve integration within schools. This is important because there is some evidence that the end of court-ordered desegregation programs have moderately increased racial segregation ([Lutz \(2011\)](#)). Other research suggests that the effects of some desegregation program on educational outcomes are modest ([Angrist and Lang \(2004\)](#)). My contribution is to show that integration at the school level does not necessarily lead to interactions among students of different groups. Furthermore, the change in the relative diversity in the school may have heterogeneous and potentially nonlinear effects on the levels of within-school segregation and welfare.

The rest of the paper is organized as follows. Section 2 briefly describes the theoretical model developed in [Mele \(2017a\)](#). Section 3 develops the estimation strategy and provides an overview of the data. Section 4 report the posterior estimates and the policy experiments. Some of the computational details are provided in Appendix.

## 2 A Model of Network Formation

In this section, I briefly present the setup of [Mele \(2017a\)](#)'s model and the equilibrium likelihood. The proofs and more theoretical results are all contained in that paper. Time is

---

<sup>7</sup>For additional structural models of network formation see [Menzel \(2015\)](#), [Sheng \(2012\)](#), [DePaula et al. \(forthcoming\)](#), [Leung \(2014b\)](#), [Leung \(2014a\)](#).

discrete and there are  $n$  players in the network. Each player is characterized by a vector of observable covariates  $X_i$ , that may contain information about gender, wealth, age, location, etc. The matrix  $X = \{X_1, X_2, \dots, X_n\}$  contains these vectors for all the players, stacked by column.

The social network of friendship nominations is represented by a  $n \times n$  adjacency matrix  $g$ , whose entries  $g_{ij}$ 's are

$$g_{ij} = \begin{cases} 1 & \text{if individual } i \text{ nominates individual } j \text{ as a friend} \\ 0 & \text{otherwise} \end{cases}$$

I follow the convention in the literature, assuming  $g_{ii} = 0$ , for any  $i$ . The network is *directed*: the existence of a link from  $i$  to  $j$  does not imply the existence of the link from  $j$  to  $i$ . This modeling choice reflects the structure of the Add Health data, where friendship nominations are not necessarily mutual. Some authors refer to this data as *perceived* networks.<sup>8</sup>

Let the *realization* of the network at time  $t$  be denoted as  $g^t$  and the *realization* of the link between  $i$  and  $j$  at time  $t$  be  $g_{ij}^t$ . The network including all the current links but  $g_{ij}^t$ , i.e.  $g^t \setminus g_{ij}^t$ , is denoted as  $g_{-ij}^t$ ; while  $g_{-i}^t$  denotes the network matrix excluding the  $i$ -th row (i.e. all the links of player  $i$ ).

The network formation process follows a stochastic best-response dynamics as in [Blume \(1993\)](#). At the beginning of each period a player  $i$  is randomly selected from the population, and he meets individual  $j$ , according to a meeting probability  $\rho(ij|g^{t-1}, X)$ . Notice that  $\rho$  may depend on the previous period network and the observable characteristics. For example, people that have many friends in common may meet with higher probability than people without common friends. Or students with similar demographics have higher probability of interaction than students with different backgrounds. An implicit assumption of the model is that the player can observe the entire network and the covariates of all the agents, before making their choice about linking.

Upon meeting agent  $j$ , player  $i$  decides whether to update his link  $g_{ij}$ . The preferences of  $i$  are defined over networks and covariates. The utility of player  $i$  from network  $g$  and covariates  $X$  is given by

$$U_i(g, X) = \underbrace{\sum_{j=1}^n g_{ij} u_{ij}}_{\text{direct friends}} + \underbrace{\sum_{j=1}^n g_{ij} g_{ji} m_{ij}}_{\text{mutual friends}} + \underbrace{\sum_{j=1}^n g_{ij} \sum_{\substack{k=1 \\ k \neq i, j}}^n g_{jk} v_{ik}}_{\text{friends of friends}} + \underbrace{\sum_{j=1}^n g_{ij} \sum_{\substack{k=1 \\ k \neq i, j}}^n g_{ki} w_{kj}}_{\text{popularity}} \quad (1)$$

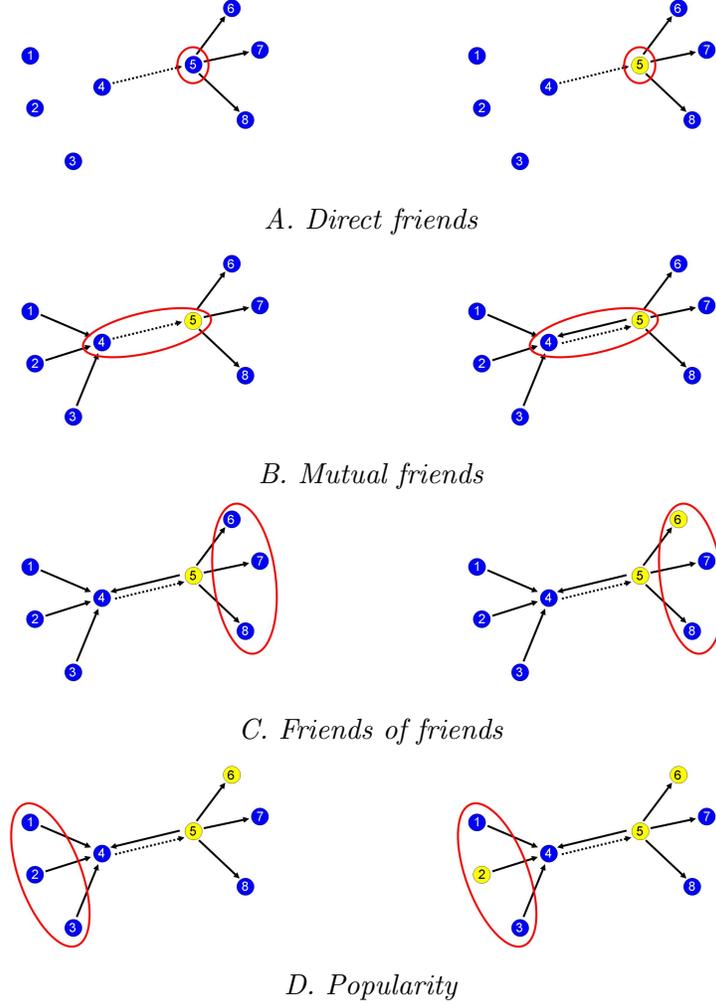
where  $u_{ij} \equiv u(X_i, X_j)$ ,  $m_{ij} \equiv m(X_i, X_j)$ ,  $v_{ij} \equiv v(X_i, X_j)$  and  $w_{ij} \equiv w(X_i, X_j)$  are (bounded) real-valued functions of the attributes. The utility of the network is the sum of the net benefits received from each link. The total benefit from an *additional link* has four components.

When a player creates a link to another individual, he receives a *direct* net benefit  $u_{ij}$ .

---

<sup>8</sup>See [Wasserman and Faust \(1994\)](#) for references.

Figure 1: Components of the utility function



The network contains  $n = 8$  agents, belonging to two groups: blue and yellow. All the panels show a situation in which player 4 decides whether to form a link to individual 5 (the dashed arrow from 4 to 5). Agent 4 receives different direct utility when he links a blue (Panel A, left) or a yellow (Panel A, right) individual. Agent 4's utility from an additional link is different if the link is unilateral (Panel B, left) or reciprocated (Panel B, right). Furthermore, agent 4's utility from friends of friends varies with their socioeconomic composition: 3 blue individuals (Panel C, left) provide different utility than 2 blue and 1 yellow agents (Panel C, right). Finally, agent 4 values how his new link affects his popularity, since he creates a new indirect friendship for those who already have a link to him (agents 1, 2 and 3). The utility of a link to agent 5 (which is yellow) when agents 1, 2 and 3 are all blue (Panel D, left) is different than when agent 2 is yellow and 1 and 3 are blue (Panel D, right).

The direct utility includes both costs and benefits and it may possibly be negative: when only homophily enters payoffs of direct links, the net utility  $u_{ij}$  is positive if  $i$  and  $j$  belong to the same group, while it is negative when they are of different types. This is illustrated

in Panel A of Figure 1 with a simple network of 8 agents. Each agent can belong to either the blue group or the yellow group. The link that agent 4 forms to individual 5 provides different direct utility in the two networks, since the identity of 5 is different: blue for the left network and yellow for the right one. In many models this component is parameterized as  $u_{ij} = b_{ij} - c_{ij}$ , where  $b_{ij}$  indicates the (gross) benefit and  $c_{ij}$  the cost of forming the additional link  $g_{ij}$ . I use the notation  $u_{ij}$ , since it does not require assumptions on the cost function.

The players receive additional utility  $m_{ij}$  if the link is mutual; friendship is valued differently if the other agent reciprocates. An agent may perceive another individual as a friend, but that person may not perceive the relationship in the same way. Panel B of Figure 1 isolates this component: a link from agent 4 to agent 5 has a different value if agent 5 reciprocates (right network).

The players value the composition of friends of friends. When  $i$  is deciding whether to befriend  $j$ , she observes  $j$ 's friends and their socioeconomic characteristics. Each of  $j$ 's friend provides additional utility  $v(X_i, X_k)$  to  $i$ . In this model, an agent who has the opportunity to form an additional link, values a white student with three Hispanic friends as a different good than a white student with two white friends and one African American friend.<sup>9</sup> In other words, individuals value both *exogenous* heterogeneity and *endogenous* heterogeneity: the former is determined by the socioeconomic characteristics of the agents, while the latter arises endogenously with the process of network formation. I assume that only friends of friends are valuable and they are perfect substitutes: individuals do not receive utility from two-links-away friends. In Panel C of Figure 1, from the perspective of agent 4, agent 5 in the left network is a different good than agent 5 in the right network, since the composition of his friends is different.

The fourth component corresponds to a *popularity effect*. Consider Panel D in Figure 1. When agent 4 forms a link to agent 5, he automatically creates an indirect link for agents 1, 2 and 3. Thus agent 4 generates an externality. For example, suppose there is homophily in indirect links. Then in the left network the externality is negative for all three agents (1, 2 and 3); and in the right network it is negative for 1 and 3, but positive for 2. Therefore, in the left network the popularity of 4 goes down, while in the right network the fall in popularity is less pronounced.

Conditional on the meeting  $m^t = ij$ , player  $i$  updates the link  $g_{ij}$  to maximize his current utility, taking the existing network  $g_{-ij}^t$  as given. I assume that the agents do not take into account the effect of their linking strategy on the future evolution of the network. The players have *complete information*, since they can observe the entire network and the individual attributes of all agents. Before updating his link to  $j$ , individual  $i$  receives an idiosyncratic shock  $\varepsilon \sim F(\varepsilon)$  to his preferences that the econometrician cannot observe. This shock models unobservables that could influence the utility of an additional link, e.g.

---

<sup>9</sup>A similar assumption is used in [De Marti and Zenou \(2009\)](#) where the agents' cost of linking depend on the racial composition of friends of friends. Their model is an extension of the connection model of [Jackson and Wolinsky \(1996\)](#), and the links are formed with mutual consent. The corresponding network is undirected.

mood, gossips, fights, etc. Player  $i$  links agent  $j$  at time  $t$  if and only if it is a best response to the current network configuration, i.e.  $g_{ij}^t = 1$  if and only if

$$U_i(g_{ij}^t = 1, g_{-ij}^{t-1}, X) + \varepsilon_{1t} \geq U_i(g_{ij}^t = 0, g_{-ij}^{t-1}, X) + \varepsilon_{0t}. \quad (2)$$

I assume that when the equality holds, the agent plays the status quo.<sup>10</sup>

**ASSUMPTIONS.** *The model satisfies the following assumptions:*

1. (*Preferences*) *The payoffs are such that  $m(X_i, X_j) = m(X_j, X_i)$  and  $w(X_k, X_j) = v(X_k, X_j)$  for all players  $i, j, k$ .*
2. (*Meetings*) *Any meeting is possible, i.e.,  $\rho(ij|g^{t-1}, X) = \rho(ij|g_{-ij}^t, X) > 0$  for any pair of players  $i, j$ .*
3. (*Shocks*) *Before deciding whether to update a link, players receive a stochastic shock that follows a Type I extreme value distribution, i.i.d. among links and across time.*

The first assumption about symmetry of  $m_{ij}$  is needed for identification: two individuals with the same exogenous characteristics  $X_i = X_j$  (say two males, whites, enrolled in eleventh grade) who form a mutual link receive the same  $u_{ij}$  and  $m_{ij}$ , but they may have different utilities from that additional link because of the composition of their friends of friends and their popularity. Therefore, this part of the assumption helps in identifying the utility from indirect links and popularity.

When  $i$  forms a link to  $j$ ,  $i$  creates an externality for all  $k$ 's who have linked her: any such  $k$  now has an additional indirect friend, i.e.  $j$ , who agent  $k$  values by an amount  $v(X_k, X_j)$ . When  $w(X_k, X_j) = v(X_k, X_j)$ , an individual  $i$  values his popularity effect as much as  $k$  values the indirect link to  $j$ , i.e.,  $i$  internalizes the externality he creates.<sup>11</sup>

The second assumption on the meeting process guarantees that any pair of agents can meet. The main implication is that any equilibrium network can be reached with positive probability. For example, a discrete uniform distribution satisfies this assumption.

Finally the third assumption allows the Markov chain to escape from the nash networks, eliminating absorbing states and making the model ergodic.

As a consequence of these assumptions, the network formation process is a potential game, where all the incentives of the players can be summarized by an aggregate function of the network  $Q$ .

$$Q(g, X) = \sum_{i=1}^n \sum_{j=1}^n g_{ij} u_{ij} + \sum_{i=1}^n \sum_{j>i}^n g_{ij} g_{ji} m_{ij} + \sum_{i=1}^n \sum_{\substack{j=1 \\ j \neq i}}^n \sum_{\substack{k=1 \\ k \neq i, j}}^n g_{ij} g_{jk} v_{ik}, \quad (3)$$

<sup>10</sup>This assumption does not affect the main result and is relevant only when the distribution of the preference shocks is discrete.

<sup>11</sup>This restriction of the preferences guarantees the model's coherency in the sense of [Tamer \(2003\)](#). In simple words, this part of the assumption guarantees that the system of conditional linking probabilities implied by the model generates a proper joint distribution of the network matrix. Similar restrictions are also encountered in spatial econometrics models ([Besag, 1974](#)) and in the literature on qualitative response models ([Heckman, 1978](#); [Amemiya, 1981](#))

The potential is such that, for any player  $i$  and any link  $g_{ij}$  we have

$$Q(g_{ij}, g_{-ij}, X) - Q(1 - g_{ij}, g_{-ij}, X) = U_i(g_{ij}, g_{-ij}, X) - U_i(1 - g_{ij}, g_{-ij}, X)$$

Consider two networks,  $g = (g_{ij}, g_{-ij})$  and  $g' = (1 - g_{ij}, g_{-ij})$ , that differ only with respect to one link,  $g_{ij}$ , chosen by individual  $i$ : the difference in utility that agent  $i$  receives from the two networks,  $U_i(g, X) - U_i(g', X)$ , is exactly equal to the difference of the *potential* function evaluated at the two networks,  $Q(g, X) - Q(g', X)$ . That is, the potential is an aggregate function that summarizes both the state of the network and the deterministic incentives of the players in each state.

The model generates a Markov Chain of networks that converges to a unique stationary distribution  $\pi$

$$\pi(g, X) = \frac{\exp[Q(g, X)]}{\sum_{\omega \in \mathcal{G}} \exp[Q(\omega, X)]}, \quad (4)$$

In the long-run the system spends more time in network states that have high potential. It can be shown that these networks correspond to Nash equilibria of a model without any shock to the preferences (Mele (2017a), Jackson and Watts (2001), Monderer and Shapley (1996)).

### 3 Estimation Strategy

To estimate the model, I assume that the payoff functions depend on a vector of parameters  $\theta = (\theta_u, \theta_m, \theta_v)$ :

$$\begin{aligned} u_{ij}(\theta_u) &= u(X_i, X_j, \theta_u) \\ m_{ij}(\theta_m) &= m(X_i, X_j, \theta_m) \\ v_{ij}(\theta_v) &= v(X_i, X_j, \theta_v) \end{aligned}$$

Assuming that the observed network is a draw from the stationary distribution of the theoretical model, we can use the distribution in (4) as likelihood of the network data. However, this imposes a computational challenge, since the likelihood depends on the normalizing constant

$$c(\mathcal{G}, X, \theta) = \sum_{\omega \in \mathcal{G}} \exp[Q(\omega, X, \theta)]. \quad (5)$$

whose exact evaluation is computationally infeasible even for very small networks. To be concrete, consider a small network with  $n = 10$  agents. From (??) we know that  $c(\mathcal{G}, X, \theta) = \sum_{\omega \in \mathcal{G}} \exp[Q(\omega, X, \theta)]$ . To compute the constant at the current parameter  $\theta$  we would need to evaluate the potential function for all  $2^{90} \simeq 10^{27}$  possible networks with 10 agents and compute their sum. This task would take a very long time even for a state-of-the-art supercomputer. In general with a network containing  $n$  players, we have to sum over

$2^{n(n-1)}$  possible network configurations.<sup>12</sup> Therefore direct evaluation of the likelihood is impossible. It is easy to show that the first and second order conditions for the maximum likelihood problem also depend on the normalizing constant. The same problem arises if we use a Bayesian approach and standard Markov Chain Monte Carlo samplers to estimate the posterior

$$p(\theta|g, X) = \frac{\pi(g, X, \theta) p(\theta)}{\int_{\Theta} \pi(g, X, \theta) p(\theta) d\theta}. \quad (6)$$

because equation (6) contains the normalizing constant in the likelihood.

### 3.1 Estimation Algorithm

To solve this challenging estimation problem, I use a variation of the *exchange algorithm*, first developed by Murray et al. (2006) for distribution with intractable normalizing constants and adapted to network models by Caimo and Friel (2011) and Mele (2017a). This algorithm uses a double Metropolis-Hastings step to avoid the computation of the normalizing constant  $c(\mathcal{G}, X, \theta)$  in the likelihood.<sup>13</sup>

While several authors have proposed similar algorithms in the related literature on Exponential Random Graphs Models (ERGM),<sup>14</sup> the models estimated with this methodology typically have very few parameters and use data from very small networks. To the best of my knowledge, this is the first attempt to estimate a high-dimensional model using data from multiple networks.

In this section I describe the algorithm for a single network, while in the appendix I provide the extension for multiple independent networks.<sup>15</sup> This is especially important for policy: schools may have unobserved differences that impact the network formation process and using multiple networks may partially correct for that.

The idea of the algorithm is to sample from an augmented distribution using an auxiliary variable. At each iteration, the algorithm proposes a new parameter vector  $\theta'$ , drawn from a suitable proposal distribution  $q_{\theta}(\theta'|\theta)$ ; in the second step, it samples a network  $g'$  (the auxiliary variable) from the likelihood  $\pi(g', X, \theta')$ ; finally, the proposed parameter is accepted

<sup>12</sup>A supercomputer that can compute  $10^{12}$  potential functions in 1 second would take almost 40 million years to compute the constant once for a network with  $n = 10$  players. The schools used in the empirical section have between 20 and 159 enrolled students. This translates into a minimum of  $2^{380}$  and a maximum of  $2^{25122}$  possible network configurations.

<sup>13</sup>This improvement comes with a cost: the algorithm may produce MCMC chains that have very poor mixing properties (Caimo and Friel, 2011) and high autocorrelation. I partially correct for this problem by carefully calibrating the proposal distribution. In this paper I use a random walk proposal. Alternatively one could update the parameters in blocks or use recent random block techniques as in Chib and Ramamurthy (2009) to improve convergence and mixing.

<sup>14</sup>Caimo and Friel (2011) use the exchange algorithm to estimate ERGM. They improve the mixing of the sampler using the snooker algorithm. Koskinen (2008) proposes the Linked Importance Sampler Auxiliary variable (LISA) algorithm, which uses importance sampling to provide an estimate of the acceptance probability. Another variation of the algorithm is used in Liang (2010).

<sup>15</sup>When the data consist of several independent school networks, I use a parallel version of the algorithm that stores each network in a different processor. Each processor runs the simulations independently and the final results are summarized in the master processor, that updates the parameters for next iteration. Details in Appendix.

with a probability  $\alpha_{ex}(\theta, \theta')$ , such that the Markov chain of parameters generated by these update rules, has the posterior (6) as unique invariant distribution.

**ALGORITHM 1. (APPROXIMATE EXCHANGE ALGORITHM)**

Fix the number of simulations  $R$ . At each iteration  $t$ , with current parameter  $\theta_t = \theta$  and network data  $g$ :

1. Propose a new parameter  $\theta'$  from a distribution  $q_\theta(\cdot|\theta)$ ,

$$\theta' \sim q_\theta(\cdot|\theta). \quad (7)$$

2. Simulate  $R$  networks from the stationary distribution of the model and collect the last simulated network  $g' \sim \mathcal{P}_{\theta'}^{(R)}(g'|g)$ , using the following steps (2.1) and (2.2) at each iteration:

- (2.1) At iteration  $r$ , with current network  $g_r$  and proposed parameter  $\theta'$ , start the simulations at network  $g$  and propose a network  $g^*$  from a proposal distribution

$$g^* \sim q_g(g^*|g_r) \quad (8)$$

- (2.2) Update the network according to

$$g_{r+1} = \begin{cases} g^* & \text{with prob. } \alpha_{mh}(g_r, g^*) \\ g_r & \text{with prob. } 1 - \alpha_{mh}(g_r, g^*) \end{cases} \quad (9)$$

where

$$\alpha_{mh}(g_r, g^*) = \min \left\{ 1, \frac{\exp [Q(g^*, X, \theta)] q_g(g_r|g^*)}{\exp [Q(g_r, X, \theta)] q_g(g^*|g_r)} \right\} \quad (10)$$

3. Update the parameter according to

$$\theta_{t+1} = \begin{cases} \theta' & \text{with prob. } \alpha_{ex}(\theta, \theta', g', g) \\ \theta & \text{with prob. } 1 - \alpha_{ex}(\theta, \theta', g', g) \end{cases}$$

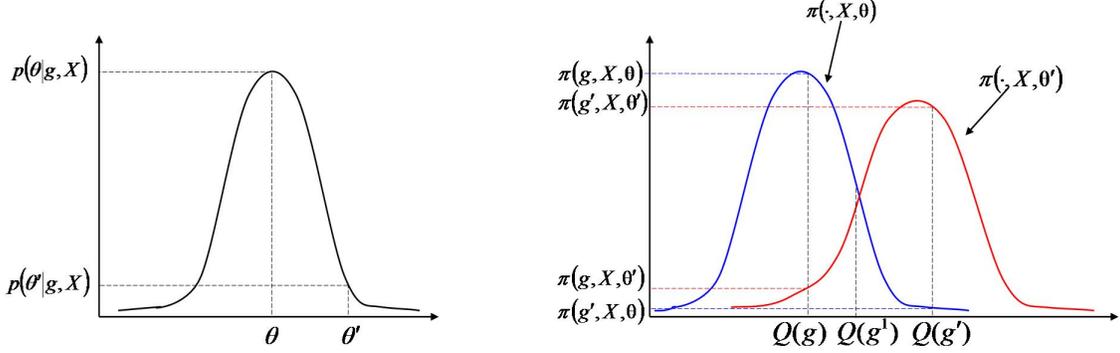
where

$$\alpha_{ex}(\theta, \theta', g', g) = \min \left\{ 1, \frac{\exp [Q(g', X, \theta)] p(\theta') q_\theta(\theta|\theta') \exp [Q(g, X, \theta')]}{\exp [Q(g, X, \theta)] p(\theta) q_\theta(\theta'|\theta) \exp [Q(g', X, \theta')]} \right\}. \quad (11)$$

The appeal of this algorithm is that all quantities in the acceptance ratio (11) can be evaluated: there are no integrals or normalizing constants to compute. I use a parallelized version of the algorithm to estimate the model using multiple school networks. Here I explain the intuition of how the sampler works, with the help of Figure 2; the algorithm's details and the proof of convergence are presented in Appendix B of Mele (2017a).

For ease of exposition, assume that the prior is relatively flat, so that  $p(\theta)/p(\theta') \simeq 1$ . Suppose we start the sampler from a parameter  $\theta$  that has high posterior probability, given

Figure 2: The Exchange Algorithm



A. Posterior Distribution

B. Two Stationary Equilibria

The graph on the left is the posterior distribution, given the data. The graph on the right represents two stationary equilibria of the model, one at parameter  $\theta$  (blue) and one at parameter  $\theta'$  (red). The iteration  $t$  starts with parameter  $\theta$ . It is proposed to update the parameter using proposal  $\theta'$ . The algorithm start sampling networks from the stationary distribution at parameter  $\theta'$  (red) and quickly moves from  $g$  to  $g'$ . The probability of accepting the proposed parameter  $\theta'$  is proportional to the ratio  $\frac{\pi(g', X, \theta)}{\pi(g', X, \theta')} \frac{\pi(g, X, \theta')}{\pi(g, X, \theta)}$ , which is small as indicated in the graph. In summary, a move from the high density region of the posterior ( $\theta$ ) to a low density region ( $\theta'$ ) is likely to be rejected. For the same reasoning a move from  $\theta'$  to  $\theta$  is very likely to be accepted. Therefore the algorithm produces samples from the correct posterior distribution.

the data  $g$ . That is, there is good agreement between the data and the parameter, so it is likely that the data are generated from a model with parameter  $\theta$ . This is displayed on the left panel of Figure 2. Now, suppose we propose a parameter  $\theta'$  that belongs to a low probability region of the posterior. This means that there is a low probability that the observed network  $g$  is generated by parameter  $\theta'$ . As a consequence the ratio

$$\frac{p(\theta'|g, X)}{p(\theta|g, X)} \simeq \frac{\pi(g, X, \theta')}{\pi(g, X, \theta)}$$

would be very small, as indicated in the right panel of Figure 2. Let's start the network simulations using parameter  $\theta'$ . The sequence of simulated networks will start approaching the new stationary distribution  $\pi(\cdot, X, \theta')$ , moving away from the stationary distribution  $\pi(\cdot, X, \theta)$ . This is indicated in Figure 2 with a simulation of 2 steps: starting from  $g$  we obtain two networks,  $g^1$  and  $g'$ . Network  $g'$  is closer to a high probability region of  $\pi(\cdot, X, \theta')$  than to a high probability region of  $\pi(\cdot, X, \theta)$ , as long as the algorithm was run for a sufficiently large number of steps  $R$ . It also follows that the ratio

$$\frac{\pi(g', X, \theta)}{\pi(g', X, \theta')} \tag{12}$$

is small. Notice that the product of the likelihood ratios does not contain any normalizing constant, that is

$$\begin{aligned} \frac{\pi(g', X, \theta)}{\pi(g', X, \theta')} \frac{\pi(g, X, \theta')}{\pi(g, X, \theta)} &= \frac{\exp [Q(g', X, \theta)] \exp [Q(g, X, \theta')]}{\exp [Q(g, X, \theta)] \exp [Q(g', X, \theta')]} \frac{c(\mathcal{G}, X, \theta')}{c(\mathcal{G}, X, \theta)} \frac{c(\mathcal{G}, X, \theta)}{c(\mathcal{G}, X, \theta')} \\ &= \frac{\exp [Q(g', X, \theta)]}{\exp [Q(g, X, \theta)]} \frac{\exp [Q(g, X, \theta')]}{\exp [Q(g', X, \theta')]} \end{aligned}$$

The product (13) is contained in the probability (11), and thus the probability of accepting the transition from  $\theta$  to the proposed parameter  $\theta'$  is small. As a consequence, the proposed parameter  $\theta'$  is very likely to be rejected. By the same reasoning, if we start the sampler at  $\theta'$  and propose an update to  $\theta$ , the transition is very likely to be accepted.

In summary, the sampler is likely to accept proposals that move towards high density regions of the posterior, but it is likely to reject proposals that move towards low density regions of the posterior. Therefore, it produces samples of parameters that closely resemble draws from the posterior distribution (6).

The formal statement about convergence and ergodicity is contained in Theorem 6 of Mele (2017a). The theorem states that the algorithm produces good samples as long as the number of simulated networks  $R$  in step 2 of estimation algorithm simulations is large enough and the posterior simulations are run for a sufficient number of iterations. The researcher can improve convergence through a careful choice of the initial network and proposal distributions. In the empirical implementation of the algorithm, I use several alternative network proposals  $q_g(\cdot|\cdot)$ . First, a move that updates only one link per iteration, proposing to swap the link value. At each iteration a random pair of agents  $(i, j)$  is selected from a discrete uniform distribution, and it is proposed to swap the value of the link  $g_{ij}$  to  $1 - g_{ij}$ . Second, to improve convergence, I allow the sampler to propose bigger moves: with a small probability  $p_{inv}$ , the sampler proposes a to invert the network matrix, i.e.  $g' = \mathbf{1} - g$ , and the proposal is accepted with probability  $\alpha_{mh}(g, g')$ .<sup>16</sup> The network simulations are started at the observed network  $g$ . There are two reasons for this choice. First, in the high density region of the posterior the observed network  $g$  must have high probability according to the model. Second, Lemma 1 in Appendix B of Mele (2017a) shows that this choice guarantees faster convergence of the approximate posterior simulation algorithm to the correct posterior.

Finally, an important tuning parameter of the algorithm is  $R$ , the number of network simulations to be performed in the second step. As  $R \rightarrow \infty$  the algorithm converges to the exact exchange algorithm of Murray et al. (2006), producing exact samples from the posterior distribution. At the same time an higher value of  $R$  would increase the computational cost and result in a higher rejection rate for the proposed parameters. I do not propose an *optimal* way to choose  $R$ , but I provide some evidence with simulated data in Appendix B, showing that there is not much difference in the estimates or convergence using different

<sup>16</sup>This move is suggested in Geyer (1992) and Snijders (2002). Snijders (2002) argues that this is particularly useful in case of a bimodal distribution. This is the case for some models in the homogeneous case, as shown in Mele (2017a).

$R$ 's. The value of  $R$  has a stronger effect on the standard deviation than on the mean of the posterior, as one would expect.

### 3.2 Identification and Practical Implementation

I assume that the utility functions  $u$ ,  $m$  and  $v$  depend *linearly* on a vector of parameters. Define  $\theta_u = (\theta_{u1}, \theta_{u2}, \dots, \theta_{uP})'$ ,  $\theta_m = (\theta_{m1}, \theta_{m2}, \dots, \theta_{mL})'$  and  $\theta_v = (\theta_{v1}, \theta_{v2}, \dots, \theta_{vS})'$ . Define the functions  $H : \mathbb{R}^A \times \mathbb{R}^A \rightarrow \mathbb{R}$ .

$$\begin{aligned} u_{ij} &= u(X_i, X_j, \theta_u) = \sum_{p=1}^P \theta_{up} H_{up}(X_i, X_j) = \theta'_u \mathbf{H}_u(X_i, X_j) \\ m_{ij} &= m(X_i, X_j, \theta_m) = \sum_{l=1}^L \theta_{ml} H_{ml}(X_i, X_j) = \theta'_m \mathbf{H}_m(X_i, X_j) \\ v_{ij} &= v(X_i, X_j, \theta_v) = \sum_{s=1}^S \theta_{vs} H_{vs}(X_i, X_j) = \theta'_v \mathbf{H}_v(X_i, X_j) \end{aligned}$$

This assumption leaves room for many interesting specifications. In particular, the functions  $H$  do not exclude interactions among different characteristics, for example interactions of race and gender of both individuals. We can consider different specifications, including different sets of variables for direct, mutual and indirect links. Interactions of individual and network-level attributes are also possible.

The main consequence of the linearity assumption is that the stationary equilibrium of the model belongs to the exponential family (Lehman (1983)) and it can be written in the form

$$\pi(g, X) = \frac{\exp[\theta' \mathbf{t}(g, X)]}{\sum_{\omega \in \mathcal{G}} \exp[\theta' \mathbf{t}(\omega, X)]}, \quad (13)$$

where  $\theta = (\theta_u, \theta_m, \theta_v)'$  is a (column) vector of parameters and  $\mathbf{t}(g, X) = (t_1(g, X), \dots, t_K(g, X))$  is a vector of sufficient statistics for the network formation model. The latter vector can contain the number of links, the number of whites-to-whites links, the number of male-to-female links and so on. Interactions between different variables are possible, e.g. the number of black-males-to-white-females links, or interactions of individual controls with school-level controls.

This likelihood is very similar to the one of exponential random graph models (Snijders (2002), Frank and Strauss (1986)). My theoretical model can be interpreted as providing the microfoundations for exponential random graphs. In this sense, we can interpret the ERGM as the stationary equilibrium of a strategic game of network formation, where myopic agents follow a stochastic best response dynamics and utilities are linear functions of the parameters.

The identification of parameters follows from the theory of exponential families (Lehman, 1983). Identification is guaranteed as long as the sufficient statistics  $t(g, X)$  are not linearly dependent, provided that the data consists of multiple independent network observations.

The linear specification allows for utility functions involving network-level controls, when estimation is performed using multiple networks. This can be achieved by a specification of the parameters such as

$$\theta_p = \theta_{p0} + \sum_{c=1}^C \theta_{pc} Z_c \quad (14)$$

where  $Z_c$  is a network-level variable. The estimation methodology presented above can be applied to this specification without any change. However, estimation of a model with unobserved heterogeneity would require significant additional computational effort (see Appendix C in Mele (2017a)).

I choose somewhat vague priors for the parameters to extract most of the information from the data. I assume independent normal priors

$$p(\theta) = \mathcal{N}(\mathbf{0}, 3\mathbf{I}_P), \quad (15)$$

where  $P$  is the number of parameters.

The proposal distribution for the posterior simulation is

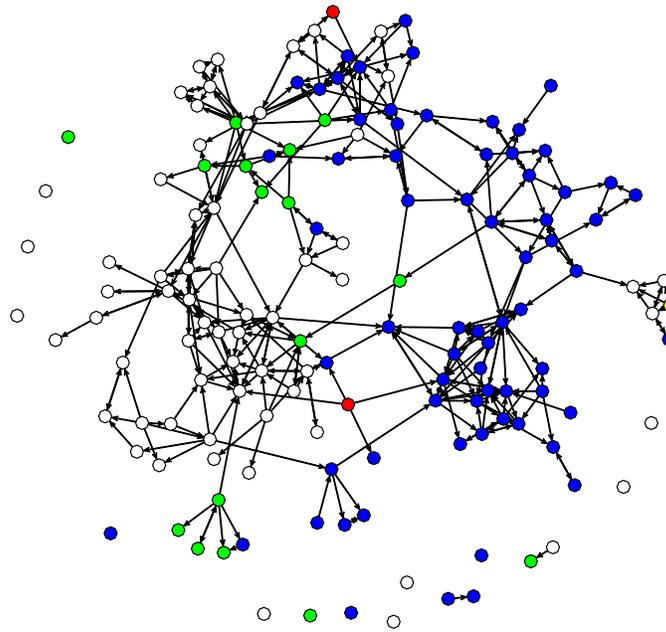
$$q_\theta(\cdot|\theta) = \mathcal{N}(\mathbf{0}, \delta\mathbf{\Sigma}), \quad (16)$$

where  $\delta$  is a scaling factor and  $\mathbf{\Sigma}$  is a covariance matrix. I use an adaptive procedure to determine a suitable  $\mathbf{\Sigma}$ . I start the iterations with  $\mathbf{\Sigma} = \lambda\mathbf{I}_P$ , where  $\lambda$  is a vector of standard deviations. I choose  $\lambda$  so that the sampler accepts at least 20%-25% of the proposed parameters, as is standard in the literature (Gelman et al., 2003; Robert and Casella, 2005). I run the chain and monitor convergence using standard methods. Once the chains have reached approximate convergence, I estimate the covariance matrix of the chains and use it as an approximate  $\mathbf{\Sigma}$  for the next set of simulations. The scaling factor is  $\delta = 2.38^2/P$  as suggested in Gelman et al. (1996).

The network sampler uses a proposal  $q_g(g|g')$ , that selects a link to be updated at each period according to a discrete uniform distribution. The probability of network inversion is  $p_{inv} = 0.01$ .

All the posterior distributions shown in the following graphs are obtained with a simulation of 100000 Metropolis-Hastings updates of the parameters. These simulations start from values found after extensive experimentation with different starting values and burn-in periods, monitoring convergence using standard methods. For each parameter update, I simulate the network for 3000 iterations to collect a sample from the stationary distribution. In Appendix, I show evidence that increasing the number of network simulations above 3000 does not change the estimates for networks of similar size as the ones used in our data.

Figure 3: A School Network



white=Whites; blue = African Americans; yellow = Asians; green = Hispanics; red = Others

Note: The graphs represent the friendship network of a school extracted from AddHealth. Each dot represents a student, each arrow is a friend nomination. The colors represent racial groups.

### 3.3 The Add Health Data

The *National Longitudinal Study of Adolescent Health* (Add Health) is a dataset containing information on a nationally representative sample of US schools. The survey started in 1994, when the 90118 participants were entering grades 7-12, and the project collected data in four successive waves.<sup>17</sup> Each student responded to an *in-school* questionnaire, and a subsample of 20745 was given an *in-home* interview to collect more detailed information about behaviors, characteristics and health status. In this paper I use only data from the *saturated sample* of Wave I, containing information on 16 schools. Each student in this sample completed both the in-school and in-home questionnaires, and the researchers made a significant effort to avoid any missing information on the students.<sup>18</sup>

I exclude the two largest schools, school 58 and 77, which have respectively 811 and 1664 students, while the third largest school has 159 students. This is to improve the speed of the computations: the estimation routine is much faster when schools are of similar size, since the parallel version of the exchange algorithm can propose a new parameter vector only after *all* the school-level simulations are done. Therefore the speed of the algorithm depends on the simulation speed of the largest school. The simulation of a school with more than 800 or 1600 students would significantly slow down the estimation. My final sample includes 1139 students enrolled in 14 schools.

The *in-school* questionnaire collects the social network of each participant. Each student was given a school roster and was asked to identify up to five male and five female friends.<sup>19</sup> I use the friendship nominations as proxy for the social network in a school. The resulting network is *directed*: Paul may nominate Jim, but this does not necessarily imply that Jim nominates Paul.<sup>20</sup> The model developed in this paper takes this feature of the data into account.

A sub-sample of 20745 students was also given an *in-home* questionnaire, that collected most of the sensible data. I use data on racial group, grade and gender of individuals. A student with a missing value in any of these variables is dropped from the sample. Each student that declares to be of Hispanic origin is considered Hispanic. The remaining non-Hispanic students are assigned to the racial group they declared. Therefore the racial categories are: White, Black, Asian, Hispanic and Other race. Other race contains Native Americans.

Additionally, I control for homophily in income. I construct the income of the family using a question from the parent questionnaire.<sup>21</sup> In the estimated models I control for income

---

<sup>17</sup>More details about the sampling design and the representativeness are contained in Moody (2001) and the Add Health website <http://www.cpc.unc.edu/projects/addhealth/projects/addhealth>

<sup>18</sup>While this sample contains no missing covariate information for the students, there are several missing values for the parental variables.

<sup>19</sup>One can think that this limit could bias the friendship data, but only 3% of the students nominated 10 friends (Moody, 2001). Moreover, the estimation routine could be easily extended to deal with missing links, as reported in Appendix.

<sup>20</sup>Some authors do not take into account this feature of the data and they recode the friendships as mutual: if a student nominates another one, the opposite nomination is also assumed.

<sup>21</sup>There are several cases in which the family income is missing. For those observations, I imputed values drawn from the unconditional income distribution of the community. An alternative but computationally very costly alternative is to introduce an additional step in the simulation, in which the imputation of missing

difference between the students and income levels.

There may be some unobservable variables that affect network formation. For example some students may be "cool" and receive more friendship links than others. To partially control for such effects, I use information from the interviewer remarks about the physical attractiveness and personality of the student interviewed. I define a dummy variable "beauty", which is equal to 1 if the interviewer told that the students was very attractive. Analogously, the dummy "personality" is equal to 1 if the interviewer responded that the personality of the student was very attractive. Additionally, I control for school fixed effects using school dummies. As an alternative to such approach, in Appendix I provide an extension of the model and estimation method that allows for unobserved heterogeneity with a significant additional computational cost.

Descriptive statistics are in Table 1. The smallest school has 20 enrolled students while the largest used in estimation has 159 students. There is a certain amount of variation in the number of links: some schools are more social and form many links per capita, while other schools have very few friendship nominations. The ratio of boys to girls is balanced in almost all schools, except school 369, where female students are large majority.

Panel A summarizes the racial composition. Many schools are almost racially homogeneous. School 1, 28, 126 and 175 are more diverse as reflected in the Racial Fragmentation index. This is an index that measure the degree of heterogeneity of a population. It is interpreted as the probability that two randomly chosen students in the school belong to different racial groups.<sup>22</sup> An index of 0 indicates that there is only one racial group and the population is perfectly homogeneous. Higher values of the index represents increasing levels of racial heterogeneity. Panel B summarizes the grade composition. Most schools offer all grades from 7th to 12th, with homogeneous population across grades. Several schools only have lower grades.

Panel C analyzes the racial and gender segregation of each school friendship network. The level of segregation is measured with the Freeman (1972) segregation index. If there is no segregation, the number of links among individuals of different groups does not depend on the group identity. The index measures the difference between the expected and actual number of links among individuals of different groups. An index of 0 means that the actual network closely resembles one in which links are formed at random. Higher values indicate more segregation. The index varies between 0 and 1, where the maximum corresponds to a network in which there are no cross-group links.

Since most schools are racially homogeneous, the measured segregation is zero. Schools with a racially diverse student population show high level of segregation for each racial group. On the other hand gender segregation is quite low and homogeneous across schools.

---

incomes is done at each iteration.

<sup>22</sup>If there are  $K$  racial groups and the share of each race is  $s_k$ , the index is

$$FRAG = 1 - \sum_{k=1}^K (s_k)^2 \tag{17}$$

Table 1: Descriptive Statistics for the schools in the Saturated Sample

School	1	2	3	7	8	28	58	77	81	88	106	115	126	175	194	369
Students	44	60	117	159	110	150	811	1664	98	90	81	20	53	52	43	52
Links	12	120	125	344	239	355	3290	3604	163	308	162	44	123	171	42	48
Females	0.5	0.517	0.419	0.44	0.5	0.587	0.473	0.483	0.531	0.522	0.531	0.55	0.491	0.538	0.512	0.654
<i>A. Racial Composition</i>																
Whites	0.5	0.95	0.983	0.981	0.973	0.42	0.978	0.055	0.98	0.989	0	1	0.472	0.769	0.977	0.942
Blacks	0.136	0	0	0.006	0.018	0.453	0.002	0.233	0	0	0.963	0	0.151	0.019	0	0
Asians	0	0	0	0	0.009	0.007	0.005	0.299	0.01	0	0	0	0.038	0.038	0	0
Hispanics	0.364	0.05	0.017	0.006	0	0.107	0.011	0.392	0.01	0	0.025	0	0.302	0.154	0.023	0.058
Others	0	0	0	0	0	0.013	0.004	0.02	0	0.011	0	0	0.038	0.019	0	0
Racial Fragn	0.599	0.095	0.034	0.037	0.053	0.606	0.044	0.699	0.04	0.022	0.072	0	0.661	0.382	0.045	0.109
<i>B. Grade Composition</i>																
7th Grade	0.159	0.2	0.128	0.145	0.227	0.173	0.002	0.001	0.112	0.144	0.506	0.4	0.491	0.462	0.488	0.538
8th Grade	0.159	0.217	0.154	0.157	0.2	0.173	0.004	0.003	0.153	0.178	0.481	0.6	0.472	0.538	0.488	0.462
9th Grade	0.114	0.2	0.12	0.214	0.136	0.2	0.289	0.004	0.153	0.122	0.012	0	0.038	0	0	0
10th Grade	0.273	0.133	0.205	0.157	0.182	0.167	0.277	0.346	0.214	0.167	0	0	0	0	0	0
11th Grade	0.136	0.167	0.179	0.164	0.118	0.14	0.223	0.345	0.265	0.211	0	0	0	0	0.023	0
12th Grade	0.159	0.083	0.214	0.164	0.136	0.147	0.205	0.301	0.102	0.178	0	0	0	0	0	0
<i>C. Segregation</i>																
Segr Whites	0	0	0	0	0	0.720	0.005	0.266	0	0	-	-	0.573	0.115	0	0
Segr Blacks	0	-	-	0	0	0.764	0	0.790	-	-	0	-	0.179	0	-	-
Segr Asian	-	-	-	-	0	0	0	0.744	0	-	-	-	0	0	-	-
Segr Hisp	0	0	0	0	-	0.429	0	0.691	-	-	0	-	0.227	0.025	0	0
Segr Other	-	-	-	-	-	0	0	0.026	-	0	-	-	0	0	-	-
Seg Gender	0.250	0.100	0.140	0.341	0.069	0.255	0.221	0.287	0.264	0.176	0.258	0.168	0.129	0.122	0.262	0.156

Descriptive statistics of the saturated sample for Wave 1 of Add Health. Segregation is measured with the [Freeman \(1972\)](#) segregation index. The index varies from 0 (perfect integration) to 1 (perfect segregation). The racial fragmentation index measures the probability that two randomly chosen students belong to different racial groups. Higher values of the index indicate a more racially heterogeneous population.

## 4 Empirical Results

### 4.1 Posterior estimates

Table 2: Posterior Distribution, Structural Model

	mean	median	std. dev.	5 pctl	95 pctl
<u>A. DIRECT UTILITY (<math>u_{ij}</math>)</u>					
CONSTANT	-4.8615	-4.8325	0.3124	-5.4380	-4.4058
SAME GENDER	0.1340	0.1366	0.1091	-0.0517	0.3062
SAME GRADE	2.0287	2.0262	0.1298	1.8141	2.2460
WHITE-WHITE	0.4408	0.4383	0.1654	0.1714	0.7288
BLACK-BLACK	0.6802	0.6886	0.1312	0.4704	0.8727
HISP-HISP	0.1649	0.1742	0.0856	0.0143	0.2904
ATTRACTIVE i (Physical)	-0.0091	-0.0100	0.1547	-0.2621	0.2546
ATTRACTIVE j (Physical)	0.1876	0.1727	0.1490	-0.0300	0.4442
ATTRACTIVE i (Personality)	0.0095	0.0093	0.1411	-0.2196	0.2426
ATTRACTIVE j (Personality)	0.2370	0.2337	0.1209	0.0401	0.4383
LOG OF (INCOME i/INCOME j)	-0.0182	-0.0172	0.0266	-0.0630	0.0230
LOG OF (INCOME i $\times$ INCOME j)	0.0685	0.0685	0.0277	0.0237	0.1147
FRACTION WHITES	-1.5015	-1.4433	0.4602	-2.3086	-0.8500
FRACTION BLACKS	1.9019	1.9234	0.1829	1.5583	2.1614
FRACTION HISP	0.2845	0.2547	0.5496	-0.5504	1.2862
WHITE-WHITE * % WHITES	-0.5351	-0.5470	0.1752	-0.7980	-0.2179
BLACK-BLACK * % BLACKS	-0.1780	-0.1227	0.3854	-0.8700	0.3626
HISP-HISP * % HISP	1.4607	1.5407	1.4443	-0.9003	3.5895
SCHOOL 1	-1.3334	-1.3416	0.4418	-2.0527	-0.6015
SCHOOL 2	1.6555	1.6527	0.2685	1.2191	2.1068
SCHOOL 3	0.3559	0.3461	0.3108	-0.1381	0.8907
SCHOOL 4	0.8271	0.8159	0.3129	0.3315	1.3665
SCHOOL 5	1.1032	1.1076	0.2803	0.6407	1.5603
SCHOOL 6	-0.7804	-0.7944	0.2720	-1.2053	-0.3189
SCHOOL 7	0.8817	0.8709	0.2863	0.4229	1.3574
SCHOOL 8	-0.7325	-0.7014	0.3790	-1.4188	-0.1569
SCHOOL 9	-0.0591	-0.1104	0.3794	-0.6177	0.6102
SCHOOL 10	2.5418	2.5386	0.3069	2.0397	3.0434
SCHOOL 11	-1.6844	-1.6501	0.3983	-2.3642	-1.0801
SCHOOL 12	1.8666	1.8578	0.2731	1.4308	2.3121
SCHOOL 13	-0.6248	-0.6084	0.3126	-1.1625	-0.1363
<u>B. MUTUAL UTILITY (<math>m_{ij}</math>)</u>					
CONSTANT	3.0074	3.0012	0.3501	2.4445	3.5969
SAME GENDER	1.2099	1.1865	0.1862	0.9457	1.5733
SAME GRADE	-1.6620	-1.6505	0.2772	-2.1371	-1.2159
WHITE-WHITE	0.1821	0.1950	0.1450	-0.0651	0.4044
BLACK-BLACK	-0.0707	-0.0591	0.2547	-0.4805	0.3512
HISP-HISP	0.7850	0.7719	0.1337	0.5703	1.0088
BOTH ATTRACTIVE (Physical)	0.1050	0.1037	0.3255	-0.4228	0.6380
BOTH NOT ATTRACTIVE (Physical)	0.0577	0.0693	0.2180	-0.3156	0.3920
BOTH ATTRACTIVE (Personality)	-0.1538	-0.1370	0.2371	-0.5641	0.2074
BOTH NOT ATTRACTIVE (Personality)	-0.1279	-0.1446	0.1762	-0.4055	0.1924
<u>C. INDIRECT UTILITY AND POPULARITY (<math>v_{ij}</math>)</u>					
CONSTANT	-0.1116	-0.1099	0.0523	-0.2002	-0.0233
SAME GENDER	-0.0992	-0.0971	0.0414	-0.1698	-0.0352
SAME GRADE	0.0224	0.0214	0.0288	-0.0246	0.0706
WHITE-WHITE	0.1535	0.1514	0.0484	0.0765	0.2343
BLACK-BLACK	0.1657	0.1613	0.0545	0.0831	0.2638
HISP-HISP	0.1673	0.1678	0.1083	-0.0141	0.3435

Estimated posterior distribution for the full structural model. The estimates are obtained with a sample of 100000 parameter simulations, and 3000 network simulations for each parameter proposal.

In Table 2, I show the estimation results for the full structural model, using the 14 schools

in the Add Health saturated sample. The table summarizes the marginal posterior distributions of the estimated parameters with posterior mean, median, standard deviation, 5th and 95th quantiles. Figures 4 and 5 display the marginal posterior distributions in blue and the posterior mean as a vertical red line.

Each estimate measures the marginal effect of the variable: for example, the parameter associated with the direct utility of WHITE-WHITE measures the marginal utility of a white student when forming a link to another white student, other things being equal.

Panel A shows the estimates for the direct utility. There is evidence of homophily in preferences: individuals prefer to form friendship links to students of the same gender, grade, race, other things being equal. Racial homophily is not homogenous across groups: African Americans show stronger preference for same race students, while the group of Hispanics has the lowest magnitude. Grade homophily is also very high.

Furthermore, there is a clear difference in how the different racial groups respond to a change in the fraction of their own group in the population. While Whites and African Americans' preference for same race students decreases with the fraction of their group in the population, the opposite occurs for Hispanics. This result is important, because it implies that different racial groups have different responses to the desegregation policies: some group may engage more in interracial friendships, while some group may segregate even more.

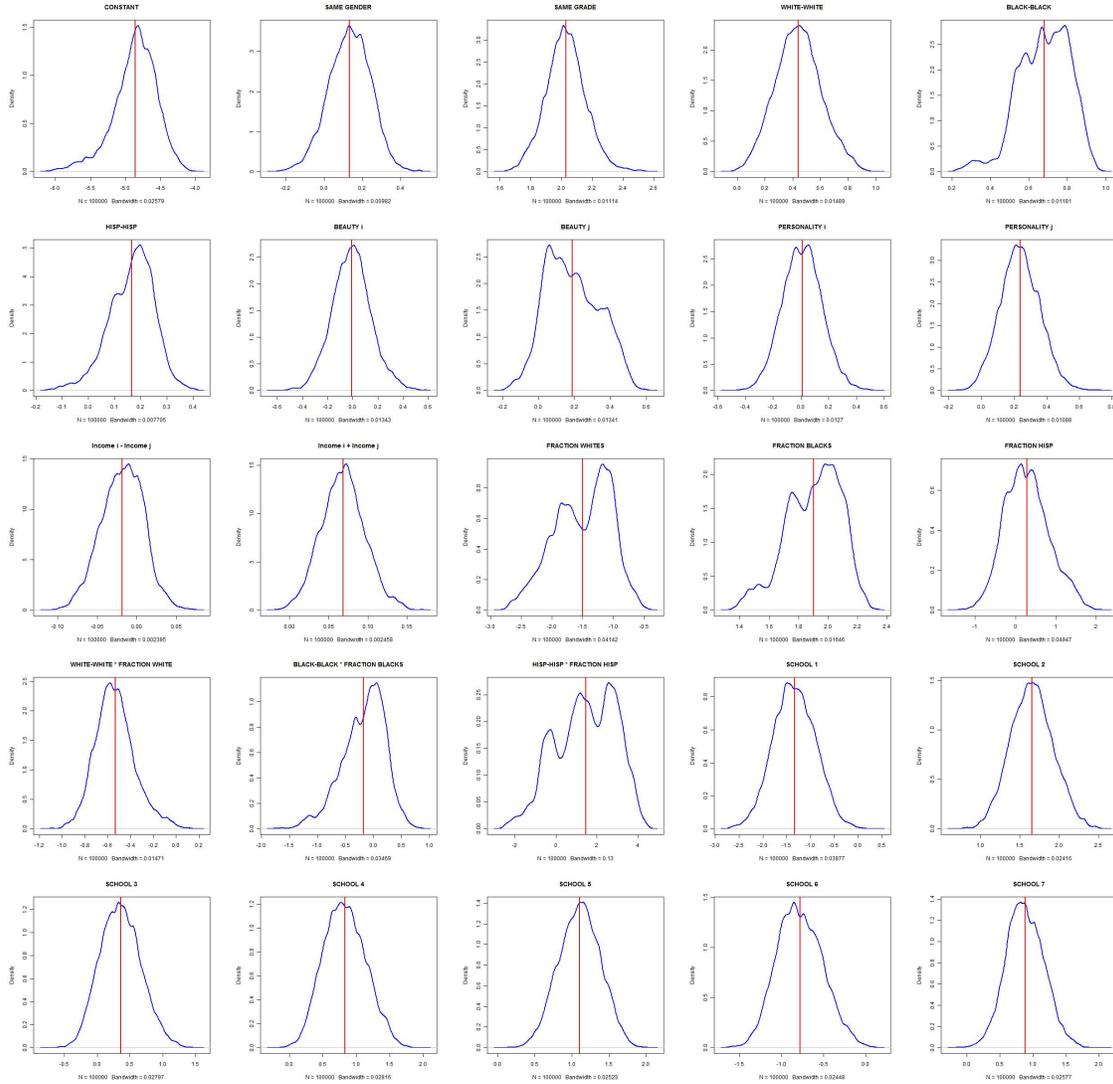
Physically attractive students have similar propensity to form friendships than the rest of the population. The same holds for students with very attractive personalities. Nonetheless, there is a propensity to form links to individuals that are physically attractive and have attractive personalities. Income differences decrease the likelihood of friendship, while higher income levels increase the number of friendships formed. The magnitude of the income effects is smaller than the effect of racial preferences. The total number of friends is higher in schools with higher fraction of minorities. The estimates for the school dummies show that there is substantial heterogeneity in the network formation process across schools, which is not accounted for by the individual characteristics.

Panel B shows the estimated parameters of the mutual utility. An additional mutual link provides positive additional utility. There is evidence of homophily in mutual links for gender, Whites and Hispanics. A mutual link to a student of the same grade decreases utility. The mutual link provides additional utility if the students have similar physical attractiveness and their personalities do not coincide.

Panel C contains estimates of the indirect and popularity effects. The negative value of the constant can be interpreted as a congestion effect: linking to students with many friends decreases utility. There is evidence of homophily in the indirect and popularity effects (except for gender), which increases the incentives of the students to segregate.

As a benchmark to evaluate policy experiments, I also estimated a model without mutual, indirect and popularity effects. The results are shown in Table 3. In this specification, the homophily effects in the direct utility are stronger than in the richer model. This is because in the full model part of the homophily effects are captured by indirect utility and popularity. Whites have much higher marginal utility of forming a link to a person of the same race than in the full model. Moreover, this specification would suggest that Whites

Figure 4: Posterior Distribution, Full Structural Model

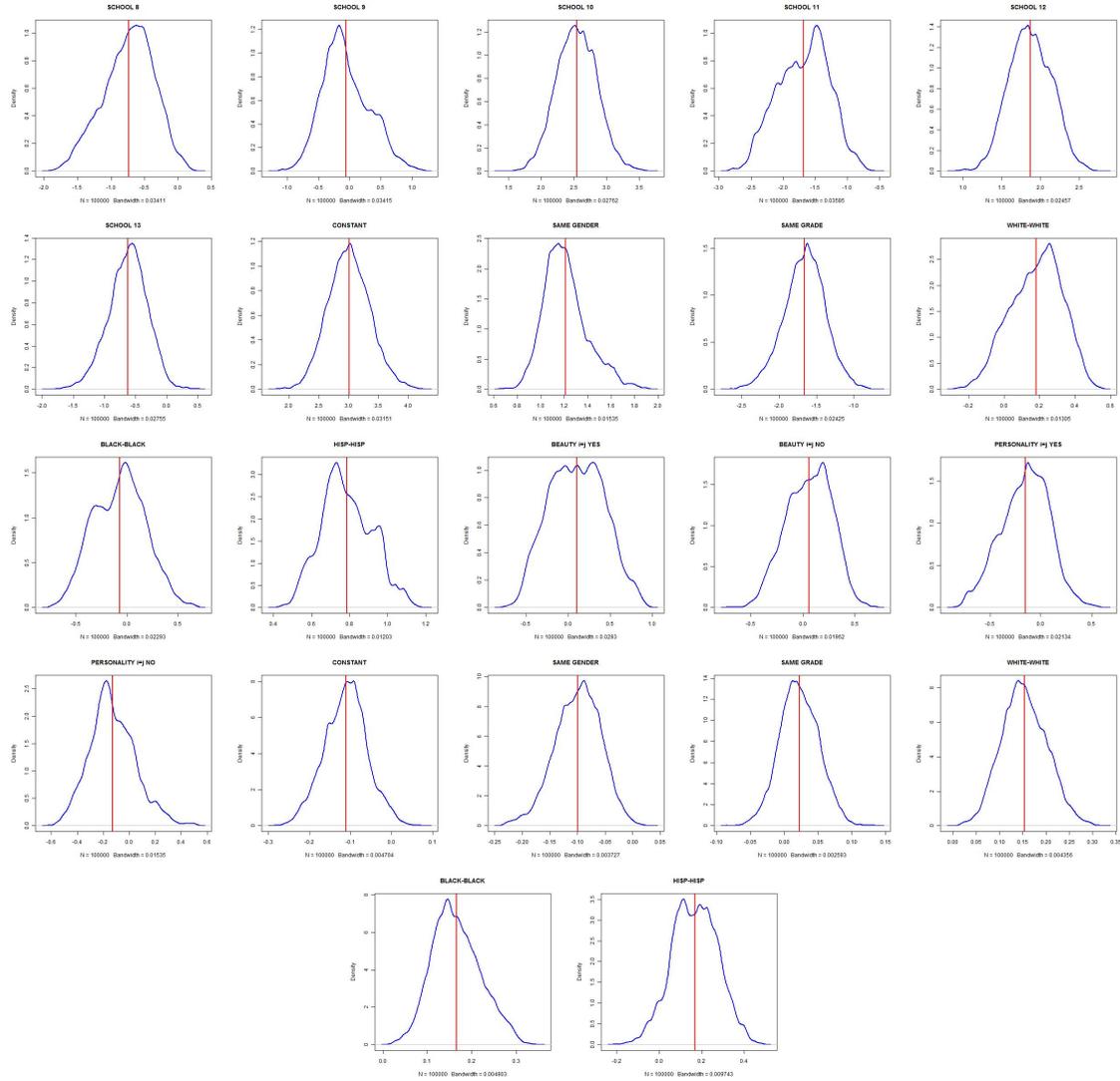


Estimated posterior distribution for the full structural model. Each graph shows the density estimate of the simulation output. The red line indicates the posterior mean. The estimates are obtained with a sample of 100000 parameter simulations and 3000 network simulations for each proposed parameter.

preferences for same race contacts are stronger than Blacks, which is the opposite result found in the full model.

An increase in the fraction of same race individuals has asymmetric effects on the racial groups: Whites and Hispanics decrease their links, while African Americans form more friendships. This is different from the full model.

Figure 5: Posterior Distribution, Full Structural Model (continued)



Estimated posterior distribution for the full structural model. Each graph shows the density estimate of the simulation output. The red line indicates the posterior mean. The estimates are obtained with a sample of 100000 parameter simulations and 3000 network simulations for each proposed parameter.

## 4.2 Policy Experiments

I use the estimated model to predict how alternative policies affect the network structure. Policy makers may be interested in pursuing policies that promote racial integration, or they may consider policies that create separate schools for boys and girls. Simulations of the model in alternative scenarios can provide a valuable benchmark on the possible effects of such policies.

Table 3: Posterior Distribution, Direct Utility only

	mean	median	std. dev.	5 pctl	95 pctl
CONSTANT	-5.3194	-5.3041	0.9443	-6.8908	-3.7122
SAME GENDER	0.4041	0.4033	0.0882	0.2610	0.5523
SAME GRADE	2.2564	2.2567	0.1048	2.0874	2.4300
WHITE-WHITE	1.6410	1.6368	0.2626	1.2165	2.0803
BLACK-BLACK	0.7567	0.7627	0.3475	0.1711	1.3160
HISP-HISP	0.9513	0.9561	0.3247	0.4061	1.4712
ATTRACTIVE i (Physical)	0.0764	0.0770	0.1401	-0.1603	0.3046
ATTRACTIVE j (Physical)	0.1713	0.1699	0.1353	-0.0461	0.4014
ATTRACTIVE i (Personality)	0.0385	0.0405	0.1403	-0.1965	0.2647
ATTRACTIVE j (Personality)	0.2914	0.2900	0.1354	0.0652	0.5131
LOG OF INCOME i/INCOME j	-0.0114	-0.0114	0.0248	-0.0521	0.0296
LOG OF INCOME i $\times$ INCOME j	0.0669	0.0663	0.0254	0.0274	0.1107
FRACTION WHITES	-1.0100	-0.9778	1.0678	-2.8420	0.6912
FRACTION BLACKS	0.5905	0.6719	0.9179	-0.9582	2.0499
FRACTION HISP	-2.6683	-2.6526	1.7131	-5.6238	0.1307
WHITE-WHITE * FRACTION WHITE	-1.4795	-1.4737	0.3316	-2.0371	-0.9475
BLACK-BLACK * FRACTION BLACKS	0.8718	0.8553	0.6620	-0.2192	1.9868
HISP-HISP * FRACTION HISP	-0.0297	0.0176	1.6081	-2.7317	2.5566
SCHOOL 1	-0.2844	-0.2649	0.6093	-1.2991	0.6866
SCHOOL 2	1.6957	1.6963	0.3413	1.1257	2.2615
SCHOOL 3	-0.1003	-0.0925	0.4302	-0.8377	0.5974
SCHOOL 4	0.5644	0.5647	0.4066	-0.1070	1.2481
SCHOOL 5	0.8732	0.8790	0.3962	0.1987	1.5249
SCHOOL 6	-0.2655	-0.2681	0.5927	-1.2511	0.7287
SCHOOL 7	0.6230	0.6206	0.3994	-0.0366	1.2829
SCHOOL 8	0.2385	0.2458	0.8298	-1.1460	1.6260
SCHOOL 9	-0.2277	-0.2259	0.5018	-1.0423	0.6200
SCHOOL 10	2.5708	2.5787	0.4515	1.8136	3.3045
SCHOOL 11	-0.6995	-0.7252	1.1728	-2.6429	1.2818
SCHOOL 12	1.4434	1.4477	0.3925	0.7884	2.0838
SCHOOL 13	-0.1609	-0.1605	0.4667	-0.9391	0.6283

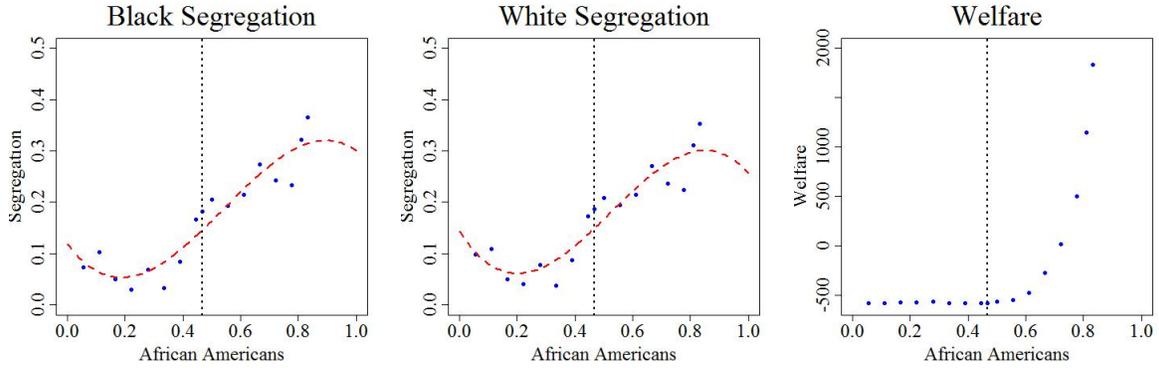
Estimated posterior distribution for the model with utility from direct links only. The estimates are obtained with a sample of 100000 parameter simulations, and 3000 network simulations for each parameter proposal.

I study the effectiveness of busing programs in promoting interracial integration. Using the posterior distribution estimated in Table 2, I simulate several busing programs that redistribute students of different racial groups among schools 88 and 106 of my sample. These are two schools with an homogeneous student population: 98.9% Whites and 96.3% African Americans, respectively. The simulated policies randomly select several (white) students from school 88 and enroll them in school 106; the same number of (black) students is randomly selected from school 106 and enrolled in school 88. This allows me to modify the ratio of Whites and African Americans in both schools and predict the levels of friendship segregation and welfare implied by the policy.<sup>23</sup>

Using 1000 draws from the estimated posterior distribution, I run the network formation model for 10000 iterations after the policy change (using ALGORITHM 1) and compute segregation and welfare of the realized networks. I use Freeman’s segregation index (see Freeman

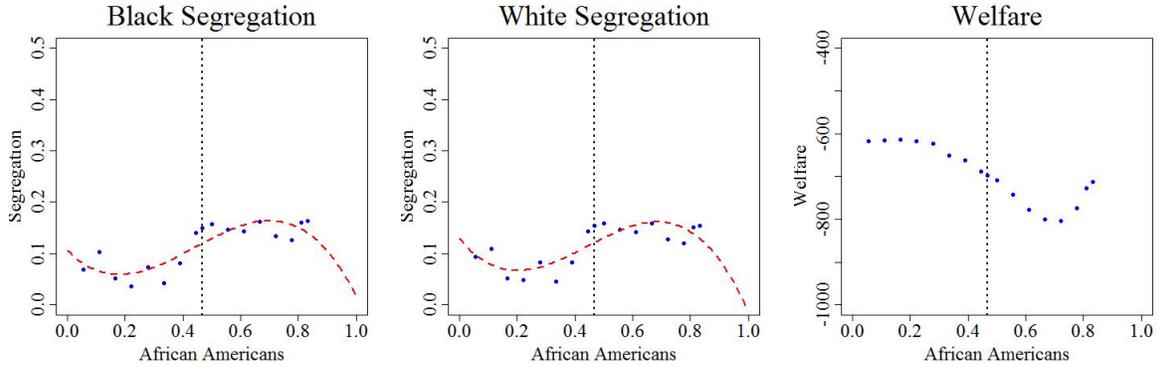
<sup>23</sup>There are several alternative ways to implement this desegregation policy. For example, one could select students based on their race, while keeping a balance for gender and income distribution. My implementation is simple. I order students of different racial groups based on their numerical ID in the data (from lowest to highest). If I need to implement a policy that re-assigns 10 students, I select the first 10 students in the list.

Figure 6: Policy Experiments for School 88 (Structural Model)



The policy experiment consists of swapping African American and White students among schools 88 and 106. The result is a change in the fraction of racial groups in the schools. The blue dots are the average segregation or welfare obtained in the simulations for each policy experiment. The vertical dotted line indicates the perfect integrated case. The red dotted line displays a cubic polynomial interpolation of the simulation results. Each simulated result was obtained with a sample of 1000 draws from the posterior distribution and a 10000 iterations of the network formation model for each posterior draw.

Figure 7: Policy Experiments for School 88 (only direct links)



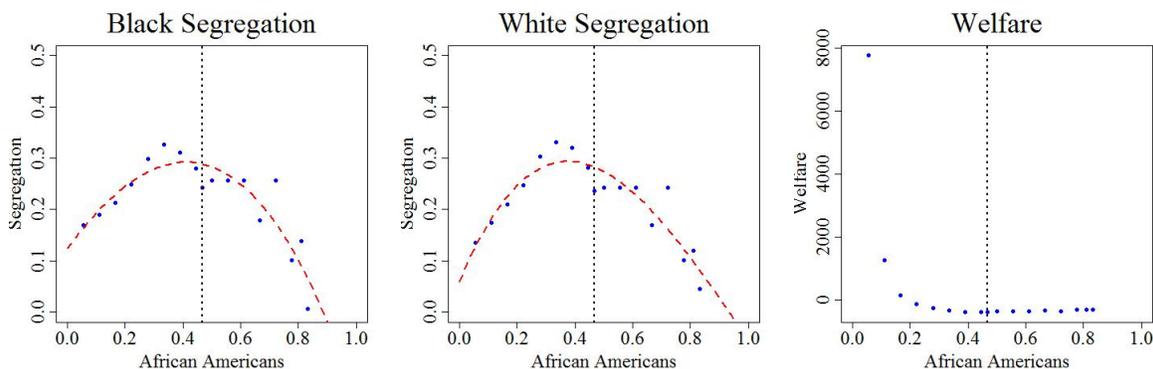
The policy experiment consists of swapping African American and White students among schools 88 and 106. The result is a change in the fraction of racial groups in the schools. The blue dots are the average segregation or welfare obtained in the simulations for each policy experiment. The vertical dotted line indicates the perfect integrated case. The red dotted line displays a cubic polynomial interpolation of the simulation results. Each simulated result was obtained with a sample of 1000 draws from the posterior distribution and a 10000 iterations of the network formation model for each posterior draw.

(1972) and Appendix B) to measure segregation for the relevant groups: Whites, African-Americans and gender. The welfare measure is the total utility of the school, i.e. the sum of all students' utilities. I simulate the model using both the full structural specification and the specification with direct utility only, to compare the policy implications. The results of

the simulations are shown in Figures 6, 7, 8 and 9. The figures report the average segregation and welfare as a function of the fraction of Whites and African Americans in the school. The blue dots are the average welfare and segregation obtained from each simulation. The vertical dotted line indicates the situation of perfect integration among schools, i.e. a policy where the students of each racial group are equally split among the schools. The red dotted line is a cubic polynomial interpolation of the simulation results.<sup>24</sup>

The simulations provide several results. First, a policy that implements perfect integra-

Figure 8: Policy Experiments for School 106 (Structural Model)



The policy experiment consists of swapping African American and White students among schools 88 and 106. The result is a change in the fraction of racial groups in the schools. The blue dots are the average segregation or welfare obtained in the simulations for each policy experiment. The vertical dotted line indicates the perfect integrated case. The red dotted line displays a cubic polynomial interpolation of the simulation results. Each simulated result was obtained with a sample of 1000 draws from the posterior distribution and a 10000 iterations of the network formation model for each posterior draw.

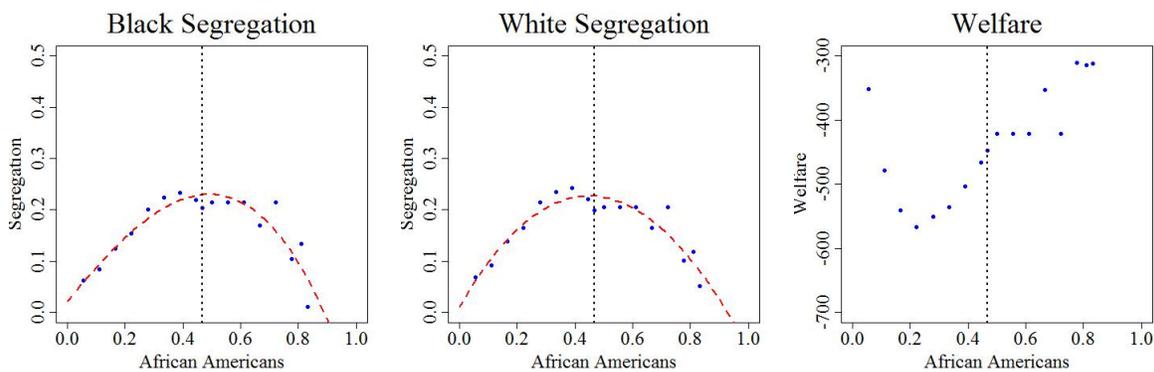
tion among the two schools is not optimal, both in terms of welfare and segregation. Perfect racial integration among the schools does not minimize expected segregation in the network of friendships and does not maximize welfare. A certain degree of segregation *among* schools is necessary to increase interracial contact and welfare *within* schools.

Second, the policy does not have the same effect on both schools. The comparison of figures 6 and 8 shows that the relationship between the fraction of each racial group and segregation is not the same. For school 88 the relationship is cubic, while for school 106 is quadratic. This further complicates the design of a good desegregation program.

Third, the model with only direct utility provides different policy recommendations than the full specification with mutual utility, indirect utility and popularity effects. This is quite evident from the graphs on welfare. For school 88, an increase of African American enrollment which does not exceed 50%, does not cause a dramatic change in welfare. The picture is completely different when looking at Figure 7. In addition, the levels of segregation predicted by the full specification are more extreme than the ones under the simpler specification.

<sup>24</sup>The curve is fitted using least squares and all the coefficients are statistically significant.

Figure 9: Policy Experiments for School 106 (only direct links)



The policy experiment consists of swapping African American and White students among schools 88 and 106. The result is a change in the fraction of racial groups in the schools. The blue dots are the average segregation or welfare obtained in the simulations for each policy experiment. The vertical dotted line indicates the perfect integrated case. The red dotted line displays a cubic polynomial interpolation of the simulation results. Each simulated result was obtained with a sample of 1000 draws from the posterior distribution and a 10000 iterations of the network formation model for each posterior draw.

## 5 Conclusions

This paper analyzed racial segregation in schools, using a structural model of network formation. The model generates segregation as an equilibrium outcome and allows me to estimate the preferences for friends belonging to the same racial group. Furthermore, the payoffs of the agents allow for homophily in indirect friendships.

I find homophily by race, both in direct and indirect links. My specification allows homophily to vary with the fraction of the racial groups in the school: an increase in the fraction of white students decreases the propensity of white students to form links within the same racial group; this is not the case for hispanics or blacks. These differences are important to understand the effect of policies that modify the relative shares of groups in the school.

To illustrate this point, I explore different desegregation policies in US schools. The model simulations provide predictions about the expected levels of segregation and welfare implied by busing programs. Perfect integration among schools could deliver unexpected results in some contexts: segregation may increase and welfare decrease. In addition, the busing program may have different effects in different schools, because of the different racial composition and the heterogeneity of preferences for same race friends. These results suggest that desegregation policies must be carefully designed to avoid unexpected outcomes.

## References

- Acemoglu, D., M. Dahleh, I. Lobel and A. Ozdaglar (2011), ‘Bayesian learning in social networks’, *Review of Economic Studies* **78**(4), 1201–1236.
- Amemiya, Takeshi (1981), ‘Qualitative response models: A survey’, *Journal of Economic Literature* **19**(4), 1483–1536.
- Ananat, Elizabeth (2011), ‘The wrong side(s) of the tracks: The causal effects of racial segregation on urban poverty and inequality’, *American Economic Journal: Applied Economics* **3**(2), 34–66.
- Angrist, Joshua D. and Kevin Lang (2004), ‘Does school integration generate peer effects? evidence from boston’s metco program’, *American Economic Review* **94**(5), 1613–1634.
- Badev, Anton (2013), Discrete games in endogenous networks: Theory and policy.
- Bandiera, Oriana and Imran Rasul (2006), ‘Social networks and technology adoption in northern mozambique’, *Economic Journal* **116**(514), 869–902.
- Besag, Julian (1974), ‘Spatial interaction and the statistical analysis of lattice systems’, *Journal of the Royal Statistical Society Series B (Methodological)* **36**(2), 192–236.
- Blume, Lawrence E. (1993), ‘The statistical mechanics of strategic interaction’, *Games and Economic Behavior* **5**(3), 387–424.
- Boucher, Vincent (2015), ‘Structural homophily’, *The International Economic Review* **56**(1), 235–264.
- Boucher, Vincent and Ismael Mourifie (forthcoming), ‘My friend far far away: A random field approach to exponential random graph models’, *Econometrics Journal*.
- Caimo, Alberto and Nial Friel (2011), ‘Bayesian inference for exponential random graph models’, *Social Networks* **33**(1), 4155.
- Chandrasekhar, Arun G. (2016), *Oxford handbook on the economics of networks.*, Oxford University Press, chapter Econometrics of network formation.
- Chib, Siddhartha and Srikanth Ramamurthy (2009), ‘Tailored random block’, *Journal of Econometrics*.
- Conley, Timothy and Christopher Udry (2010), ‘Learning about a new technology: Pineapple in ghana’, *American Economic Review* **100**(1), 35–69.
- Cooley, Jane (2010), Desegregation and the achievement gap: Do diverse peers help? working paper.

- Currarini, Sergio, Matthew O. Jackson and Paolo Pin (2009), ‘An economic model of friendship: Homophily, minorities, and segregation’, *Econometrica* **77**(4), 1003–1045.
- Currarini, Sergio, Matthew O. Jackson and Paolo Pin (2010), ‘Identifying the roles of race-based choice and chance in high school friendship network formation’, *the Proceedings of the National Academy of Sciences* **107**(11), 4857–4861.
- Cutler, David M. and Edward L. Glaeser (1997), ‘Are ghettos good or bad?’, *Quarterly Journal of Economics* **112**(3), 827–872.
- De Giorgi, Giacomo, Michele Pellizzari and Silvia Redaelli (2010), ‘Identification of social interactions through partially overlapping peer groups’, *American Economic Journal: Applied Economics* **2**(2).
- De Marti, Joan and Yves Zenou (2009), Ethnic identity and social distance in friendship formation, Cepr discussion papers, C.E.P.R. Discussion Papers.
- DePaula, Aureo (forthcoming), ‘Econometrics of network models’, *Advances in Economics and Econometrics: Theory and Applications* .
- DePaula, Aureo, Seth Richards-Shubik and Elie Tamer (forthcoming), ‘Identifying preferences in networks with bounded degree’, *Econometrica* .
- Dzemeski, Andreas (2017), An empirical model of dyadic link formation in a network with unobserved heterogeneity. Working Paper.
- Echenique, Federico and Roland Fryer (2007), ‘A measure of segregation based on social interactions’, *Quarterly Journal of Economics* **122**(2), 441–485.
- Echenique, Federico, Roland G. Fryer and Alex Kaufman (2006), ‘Is school segregation good or bad?’, *American Economic Review* **96**(2), 265–269.
- Ferrara, Eliana La and Angelo Mele (2011), Racial segregation and public school expenditure. Unpublished manuscript, IGIER Bocconi University.
- Frank, Ove and David Strauss (1986), ‘Markov graphs’, *Journal of the American Statistical Association* **81**, 832–842.
- Freeman, L. (1972), ‘Segregation in social networks’, *Sociological Methods and Research* **6**, 411–427.
- Gelman, A., G. O. Roberts and W. R. Gilks (1996), ‘Efficient metropolis jumping rules’, *Bayesian Statistics* **5**, 599–608.
- Gelman, A., J. Carlin, H. Stern and D. Rubin (2003), *Bayesian Data Analysis, Second Edition*, Chapman & Hall/CRC.

- Geyer, Charles J. (1992), ‘Practical markov chain monte carlo’, *Statistical Science* **7**, 473–511.
- Golub, Benjamin and Matthew Jackson (2011), ‘Network structure and the speed of learning: Measuring homophily based on its consequences’, *Annals of Economics and Statistics* .
- Graham, Bryan (2017), ‘An empirical model of network formation: with degree heterogeneity’, *Econometrica* **85**(4), 1033–1063.
- Heckman, James J. (1978), ‘Dummy endogenous variables in a simultaneous equation system’, *Econometrica* **46**(4), 931–959.
- Jackson, Matthew and Alison Watts (2001), ‘The existence of pairwise stable networks’, *Seoul Journal of Economics* **14**(3), 299–321.
- Jackson, Matthew and Asher Wolinsky (1996), ‘A strategic model of social and economic networks’, *Journal of Economic Theory* **71**(1), 44–74.
- Jackson, Matthew O. (2008), *Social and Economics Networks*, Princeton.
- Koskinen, Johan H. (2008), The linked importance sampler auxiliary variable metropolis hastings algorithm for distributions with intractable normalising constants. MelNet Social Networks Laboratory Technical Report 08-01, Department of Psychology, School of Behavioural Science, University of Melbourne, Australia.
- Laschever, Ron (2009), The doughboys network: Social interactions and labor market outcomes of world war i veterans. working paper.
- Lehman, E. L. (1983), *Theory of Point Estimation*, Wiley and Sons.
- Leung, Michael (2014a), A random-field approach to inference in large models of network formation. working paper.
- Leung, Michael (2014b), Two-step estimation of network-formation models with incomplete information. working paper.
- Liang, Faming (2010), ‘A double metropolis-hastings sampler for spatial models with intractable normalizing constants’, *Journal of Statistical Computing and Simulation* **80**, 1007–1022.
- Liang, Faming, Chuanhai Liu and Raymond J. Carroll (2010), *Advanced Markov Chain Monte Carlo Methods: Learning from Past Samples*, John Wiley and Sons, Ltd.
- Lutz, Byron (2011), ‘The end of court-ordered desegregation’, *American Economic Journal: Economic Policy* **3**, 130–168.
- Mele, Angelo (2017a), ‘A structural model of dense network formation’, *Econometrica* **85**, 825–850.

- Mele, Angelo (2017b), A structural model of homophily and clustering in social networks. working paper.
- Menzel, Konrad (2015), Strategic network formation with many agents, Working papers, NYU.
- Monderer, Dov and Lloyd Shapley (1996), ‘Potential games’, *Games and Economic Behavior* **14**, 124–143.
- Moody, James (2001), ‘Race, school integration, and friendship segregation in america’, *American Journal of Sociology* **103**(7), 679–716.
- Murray, Iain A., Zoubin Ghahramani and David J. C. MacKay (2006), ‘Mcmc for doubly-intractable distributions’, *Uncertainty in Artificial Intelligence* .
- Nakajima, Ryo (2007), ‘Measuring peer effects on youth smoking behavior’, *Review of Economic Studies* **74**(3), 897–935.
- Robert, Christian P. and George Casella (2005), *Monte Carlo Statistical Methods (Springer Texts in Statistics)*, Springer-Verlag New York, Inc., Secaucus, NJ, USA.
- Sheng, Shuyang (2012), Identification and estimation of network formation games.
- Snijders, Tom A.B (2002), ‘Markov chain monte carlo estimation of exponential random graph models’, *Journal of Social Structure* **3**(2).
- Tamer, Elie (2003), ‘Incomplete simultaneous discrete response model with multiple equilibria’, *The Review of Economic Studies* **70**(1), 147–165.
- Topa, Giorgio (2001), ‘Social interactions, local spillovers and unemployment’, *Review of Economic Studies* **68**(2), 261–295.
- Wasserman, Stanley and Katherine Faust (1994), *Social Network Analysis: Methods and Applications*, Cambridge University Press.

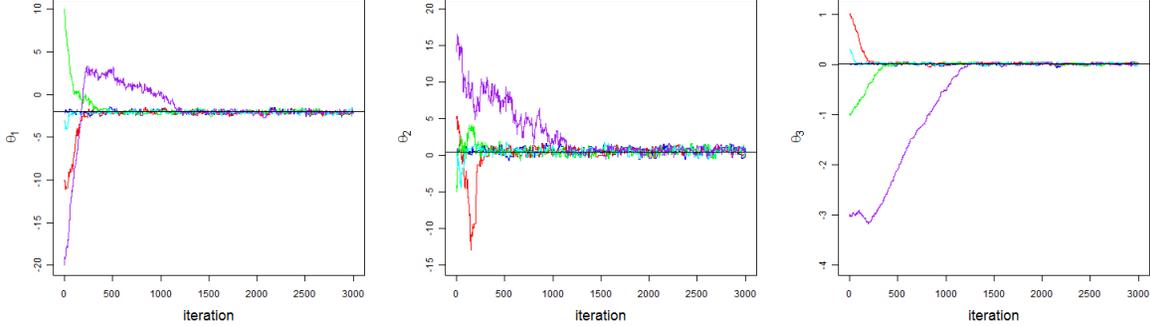
## A Computational Details

### A.1 Convergence Experiments

In this section, I provide an overview of the convergence properties of the algorithm using examples with artificial data. Assume a toy model with three parameters, with an utility function of the following form

$$U_i(g, X) = \sum_{j=1}^n g_{ij}\theta_1 + \sum_{j=1}^n g_{ij}g_{ji}\theta_2 + \sum_{j=1}^n g_{ij} \sum_{k \neq i, j; k=1}^n g_{jk}\theta_3 + \sum_{j=1}^n g_{ij} \sum_{k \neq i, j; k=1}^n g_{ki}\theta_3 \quad (18)$$

Figure 10: Convergence to the high density posterior region



Each graph shows convergence to the high density region of the posterior distribution. The curves with different colors represent chains started at overdispersed initial values. The solid black line represent the parameter that generated the data. Convergence is very fast and we can use the initial 2000 iterations as burn-in. In this example the network has  $n = 50$  agents and the number of network simulations per proposal is  $R = 3000$ .

The artificial data are generated using the vector of parameters

$$\theta = (-2.0, 0.5, 0.01) \tag{19}$$

To obtain the network dataset for the estimation, the network simulation algorithm is started at a random network and then ran for 1 million iterations. The initial random network is generated by assuming each link is independent and the probability of a link is  $p = .2$ . The last iteration of this long simulation is used as dataset in all the estimation exercises below. I report results for networks with  $n = 50$  and  $n = 100$  agents.

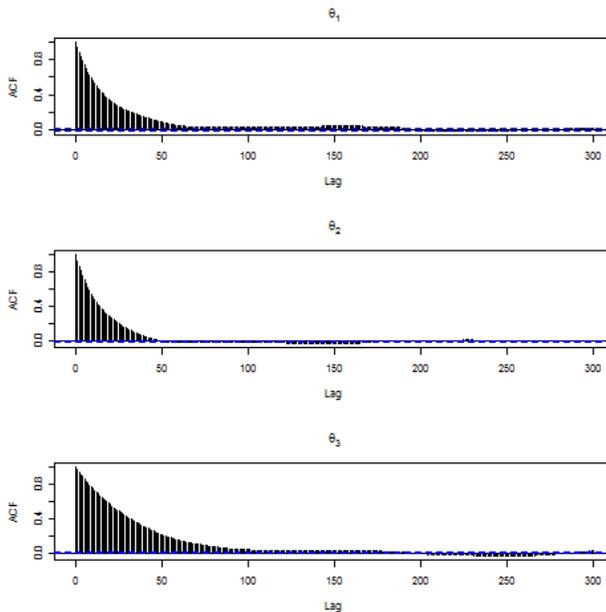
To check if the exchange algorithm converges to the correct region of the parameter space, the parameter simulations are started from 5 over-dispersed starting values

$$\begin{aligned} \theta^1 &= (-2.0, 0.5, 0.01) \\ \theta^2 &= (-10.0, 5.0, 1.0) \\ \theta^3 &= (10.0, -5.0, -1.0) \\ \theta^4 &= (-3.0, -0.05, 0.3) \\ \theta^5 &= (-20.0, 15.0, -0.3) \end{aligned}$$

In Figure 10, I display the convergence of the simulations to the high density region of the posterior. In this example the number of network simulations for each parameter proposal is  $R = 3000$ .<sup>25</sup> The solid horizontal black line represents the parameter that generated the data. Each color represents a simulation started at one of the initial values above. After 2000 iterations all the chains have reached approximate convergence to the region of the posterior that contains the data generating parameters. In Figure 11, I show the autocorrelation

<sup>25</sup>Similar results hold for different  $R$  values.

Figure 11: Convergence, autocorrelation functions



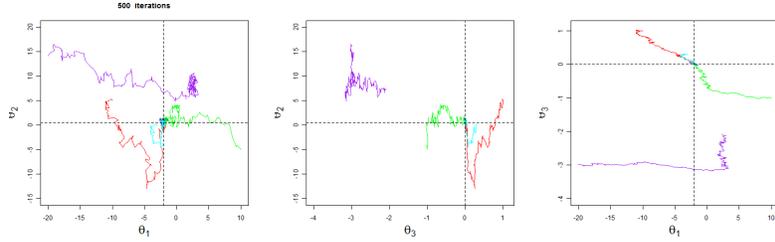
Each graph is the autocorrelation function of the chains generated by the exchange algorithm.

functions for the same example. In this example the autocorrelation disappears after 200 lags. This is mainly due to the small amount of parameters in this toy model. High-dimensional models show more persistent autocorrelation of the chains. As a consequence, the length of the simulations for the empirical application is much longer. In Figure 12 I show the same convergence properties of Figure 10 by plotting two parameters in each graph. I show 3 snapshots of the simulations: at 500, 1000 and 2000 iterations. The dashed lines intersect at the parameter values that generated the data. After 500 iterations (Panel A) almost all chains have converged to the high density region. The purple chain converges after 2000 iterations: this is because this chain corresponds to the 5th starting value, which is the quite far from the parameter that generated the network. Table 4 reports the result of estimations using different network simulation lengths, with a network of  $n = 50$  players. The table suggests that  $R = 1000$  is maybe too small, while there is not much difference among the remaining estimation results. In Table 5, I show similar results for a network with  $n = 100$  players and starting the simulations at a parameter vector  $(-20.0, 15.0, -0.3)$ .<sup>26</sup> For this network size,  $R = 3000$  would sufficient. This is the amount of simulation used in the empirical application.

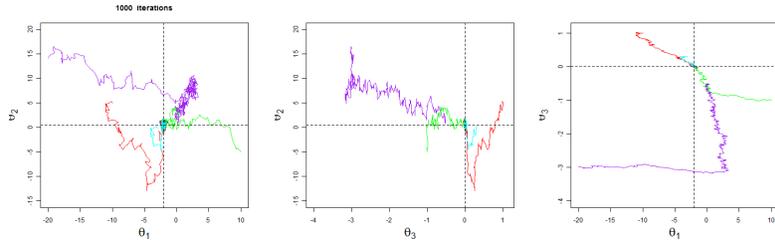
In summary, convergence in this toy model is quite fast. For high-dimensional models convergence is slower, but reasonable, in the order of 50 or 100 thousands iterations. One

<sup>26</sup>Similar results hold for alternative starting values.

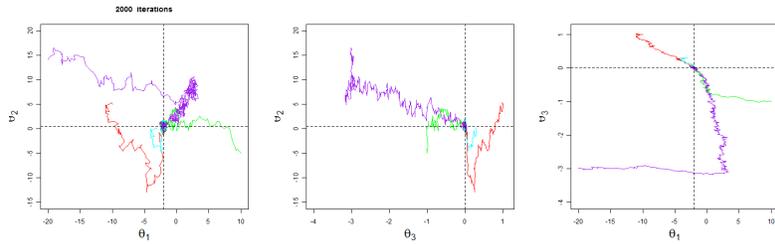
Figure 12: Convergence of the simulations



Panel A. 500 iterations



Panel B. 1000 iterations



Panel C. 2000 iterations

Three snapshots of the simulations at 500, 1000 and 2000 iterations of the fast exchange algorithm. The true parameter value is indicated by the intersection of the dashed lines. After 500 iterations only few chains have converged close to the true parameters. After 1000 the remaining chains have almost reached the high density region of the posterior. At 2000 iterations the algorithm has reached approximate convergence for all the chains.

possible strategy is to use a small  $R$  for the initial simulations: when the chain reaches approximate convergence we can increase the number of network simulations and estimate the posterior with higher precision.

## A.2 Parallel estimation with multiple networks

When data from multiple independent networks are available the estimation routines are easily adapted. Assume the researcher has data from  $C$  networks: let  $g_c$  and  $X_c$  denote the network matrix and the individual controls for network  $c$ ,  $c = 1, \dots, C$ . The aggregate data are denoted as  $g = \{g_1, \dots, g_c\}$  and  $X = \{X_1, \dots, X_c\}$ .

Table 4: Convergence Experiments,  $n = 50$

Starting value: (-2.00,0.50,0.01)						
	true		R=1000	R=2000	R=3000	R=5000
$\theta_1$	-2.000	mean	-2.0165	-2.0643	-2.077	-2.0838
		s.d.	0.2629	0.2018	0.1845	0.1635
		mc s.e.	0.0125	0.0069	0.0063	0.0051
$\theta_2$	0.500	mean	0.5387	0.6083	0.6207	0.6158
		s.d.	0.5519	0.4435	0.4144	0.4076
		mc s.e.	0.0338	0.0294	0.0189	0.0279
$\theta_3$	0.010	mean	0.0043	0.0121	0.0147	0.0175
		s.d.	0.0262	0.0201	0.0187	0.0165
		mc s.e.	0.0002	0.0001	0.0001	0.0001
Starting value: (-10.0,5.0,1.0)						
	true		R=1000	R=2000	R=3000	R=5000
$\theta_1$	-2.000	mean	-2.0131	-2.0651	-2.0688	-2.0673
		s.d.	0.2643	0.2013	0.1814	0.1655
		mc s.e.	0.0137	0.0067	0.0057	0.0046
$\theta_2$	0.500	mean	0.5542	0.6181	0.6149	0.6571
		s.d.	0.5506	0.4425	0.4228	0.4046
		mc s.e.	0.0363	0.0279	0.029	0.022
$\theta_3$	0.010	mean	0.0041	0.0119	0.0143	0.0157
		s.d.	0.0267	0.0201	0.0185	0.0167
		mc s.e.	0.0002	0.0001	0.0001	0.0001
Starting value: (10.0,-5.0,-1.0)						
	true		R=1000	R=2000	R=3000	R=5000
$\theta_1$	-2.000	mean	-2.0287	-2.0583	-2.0656	-2.0686
		s.d.	0.2548	0.2072	0.1883	0.164
		mc s.e.	0.0099	0.0081	0.0085	0.0043
$\theta_2$	0.500	mean	0.5723	0.6028	0.6275	0.6593
		s.d.	0.5418	0.4473	0.4084	0.3844
		mc s.e.	0.034	0.0224	0.0283	0.0207
$\theta_3$	0.010	mean	0.0058	0.0113	0.0128	0.016
		s.d.	0.0255	0.0211	0.0203	0.0167
		mc s.e.	0.0002	0.0001	0.0001	0.0001
Starting value: (-3.0,-0.05,0.3)						
	true		R=1000	R=2000	R=3000	R=5000
$\theta_1$	-2.000	mean	-2.016	-2.0727	-2.0884	-2.0724
		s.d.	0.2574	0.2033	0.1842	0.1625
		mc s.e.	0.01	0.0064	0.007	0.0051
$\theta_2$	0.500	mean	0.5612	0.5993	0.6354	0.6576
		s.d.	0.5436	0.4442	0.4163	0.4044
		mc s.e.	0.0346	0.027	0.0252	0.0256
$\theta_3$	0.010	mean	0.0047	0.0128	0.0158	0.0162
		s.d.	0.0254	0.0205	0.0181	0.0165
		mc s.e.	0.0002	0.0001	0.0001	0.0001

Convergence experiments using artificial data. The network contains  $n = 50$  players and the data are generated by 1 million iterations of the simulation algorithm using the true parameters. The posterior is estimated using different lengths  $R$  of the network simulation algorithm. The table reports the estimated posterior mean, standard deviation and Monte Carlo standard error for the posterior mean.

Assuming each network is drawn from the stationary equilibrium of the model, each network has distribution

$$\pi(g_c, X_c, \theta) = \frac{\exp [Q(g_c, X_c, \theta)]}{\sum_{\omega \in \mathcal{G}_c} \exp [Q(\omega_c, X_c, \theta)]} \quad (20)$$

Table 5: Convergence Experiments,  $n = 100$

		Starting value: (-20.0,15.0,-0.3)							
	true		R=1000	R=3000	R=5000	R=10000	R=30000	R=50000	R=100000
$\theta_1$	-2.000	mean	-2.0783	-2.0989	-2.0892	-2.0789	-2.0938	-2.0874	-2.1055
		s.d.	0.2613	0.1375	0.1101	0.091	0.0632	0.0703	0.0823
		mcse	0.0111	0.0015	0.0031	0.0025	0.0018	0.0016	0.0026
$\theta_2$	0.500	mean	0.341	0.5835	0.5627	0.5909	0.6475	0.6396	0.6201
		s.d.	0.7119	0.3499	0.343	0.2751	0.2284	0.2612	0.3127
		mcse	0.081	0.0176	0.0335	0.028	0.0253	0.0436	0.0364
$\theta_3$	0.010	mean	0.006	0.0096	0.0105	0.0111	0.0113	0.0124	0.0119
		s.d.	0.0114	0.0065	0.0051	0.0042	0.0032	0.0031	0.0031
		mcse	0.000020	0.000005	0.000005	0.000007	0.000003	0.000005	0.000005

Convergence experiments using artificial data. The network contains  $n = 100$  players and the data are generated by 1 million iterations of the simulation algorithm using the true parameters. The posterior is estimated using different lengths  $R$  of the network simulation algorithm. The table reports the estimated posterior mean, standard deviation and Monte Carlo standard error for the posterior mean.

Since each network is independent, the likelihood of the data  $(g, X)$  can be written as

$$\begin{aligned} \pi(g, X, \theta) &= \prod_{c=1}^C \pi(g_c, X_c, \theta) = \prod_{c=1}^C \left\{ \frac{\exp [Q(g_c, X_c, \theta)]}{c(\mathcal{G}_c, X_c, \theta)} \right\} \\ &= \frac{\exp \left[ \sum_{c=1}^C Q(g_c, X_c, \theta) \right]}{\prod_{c=1}^C c(\mathcal{G}_c, X_c, \theta)} = \frac{\exp \left[ \sum_{c=1}^C Q(g_c, X_c, \theta) \right]}{C(\mathcal{G}, X, \theta)} \end{aligned}$$

where  $\mathcal{G} = \bigcup_{c=1}^C \mathcal{G}_c$  and  $X = \{X_1, \dots, X_C\}$ . The likelihood for multiple independent networks is of the same form as the likelihood for one network observation. The structure of this likelihood makes parallelization extremely easy: each network can be simulated independently using the network simulation algorithm; at the end of the simulation we collect the last network and compute the potential; then we compute the sum of potentials and use it to compute the probability of update.

Therefore, the algorithm is modified as follows

**ALGORITHM 2. (Parallel FAST EXCHANGE ALGORITHM)**

Fix the number of simulations  $R$ . Store each network data  $(g_c, X_c)$  in a different processor/core. At each iteration  $t$ , with current parameter  $\theta_t = \theta$  and network data  $g$

1. Propose a new parameter  $\theta'$  from a distribution  $q_\theta(\cdot|\theta)$

$$\theta' \sim q_\theta(\cdot|\theta) \tag{21}$$

2. For each processor  $c$ , start the network sampler at the observed network  $g_c$ , iterating for  $R$  steps using parameter  $\theta'$  and collect the last simulated network  $g'_c$

$$g'_c \sim \mathcal{P}_{\theta'}^{(R)}(g'_c|g_c) \tag{22}$$

3. Update the parameter according to

$$\theta_{t+1} = \begin{cases} \theta' & \text{with prob. } \alpha_{pex}(\theta, \theta') \\ \theta & \text{with prob. } 1 - \alpha_{pex}(\theta, \theta') \end{cases}$$

where

$$\alpha_{pex}(\theta, \theta') = \min \left\{ 1, \frac{\exp \left[ \sum_{c=1}^C Q(g'_c, X_c, \theta) \right] p(\theta') q_\theta(\theta|\theta') \exp \left[ \sum_{c=1}^C Q(g_c, X_c, \theta') \right]}{\exp \left[ \sum_{c=1}^C Q(g_c, X_c, \theta) \right] p(\theta) q_\theta(\theta'|\theta) \exp \left[ \sum_{c=1}^C Q(g'_c, X_c, \theta) \right]} \right\} \quad (23)$$

The speed of the algorithm depends on the largest network in the data. Since each parameter update requires the result of each processor simulation there is some idle time, since small networks are simulated much faster. However, one could easily modify the algorithm to have different number of network simulations for networks of different sizes, so for each  $c$  we would have a different  $R_c$

### A.3 Freeman Segregation Index

The Freeman segregation index measures the degree of segregation in a population with two groups (Freeman, 1972). Assume there are two groups, A and B. Let  $n_{AB}$  be the total number of links that individuals of group A form to individuals of group B. Let  $n_{BA}$ ,  $n_{BB}$  and  $n_{AA}$  be analogously defined. The original index developed by Freeman (1972) is defined as

$$FSI = \frac{\mathbb{E}[n_{AB}] + \mathbb{E}[n_{BA}] - (n_{AB} + n_{BA})}{\mathbb{E}[n_{AB}] + \mathbb{E}[n_{BA}]} \quad (24)$$

When the link formation does not depend on the identity of individuals, then the links should be randomly distributed with respect to identity. Therefore, the index measures the difference between the expected and actual number of links among individuals of different groups, as a fraction of the expected links. An index of 0 means that the actual network closely resembles one in which links are formed at random. Higher values indicate more segregation. In this paper segregation is measured using the index<sup>27</sup>

$$SEG = \max\{0, FSI\} \quad (25)$$

The index varies between 0 and 1, where the maximum corresponds to a network in which there are no cross-group links.

To complete the derivation of the index, the expected number of cross-group links is computed as

$$\begin{aligned} \mathbb{E}[n_{AB}] &= \frac{(n_{AA} + n_{AB})(n_{AB} + n_{BB})}{n_{AA} + n_{AB} + n_{BA} + n_{BB}} \\ \mathbb{E}[n_{BA}] &= \frac{(n_{BA} + n_{BB})(n_{AA} + n_{BA})}{n_{AA} + n_{AB} + n_{BA} + n_{BB}} \end{aligned}$$

<sup>27</sup>The index (24) varies between -1 and 1. However, the interpretation of the index when it assumes negative values is not clear. Therefore Freeman (1972) suggests to use only when it is nonnegative, to measure the presence of segregation