

Labor Economics, 14.661. Lecture 5: Peer Effects over Networks

Daron Acemoglu

MIT

November 10, 2016.

Beyond Random Assignment

- Random assignment will not be feasible in many settings.
- This raises a host of interesting issues — specifically because in the presence of externalities, behavior will be interdependent across individuals.
- Even if there are quasi-experimental variation, how individuals will respond to this might be different, and might be a function of their expectations about how others are behaving.
- To discuss these issues and modeling and econometric approaches to tackle them, I will now consider a more general setting in which there could be a richer set of interactions.
- The modeling of this will be easier using the language of “networks”.

Network Representation

- Consider the following reduced-form model in an economy consisting of n individuals/workers.
- A (nonnegative) matrix \mathbf{G} represents the network of interactions.
 - This is also called the *adjacency matrix* or the interaction matrix—it represents who is whose “neighbor”.
- We take the network to be *directed* and *weighted*, so that the entries of \mathbf{G} are nonnegative numbers (normalized to be less than 1), and \mathbf{G} is not necessarily symmetric.
- For this lecture, let us adopt the convention that *the diagonal elements of G are zero*, and we denote the entries by g_{ij} and the i th column of \mathbf{G} by \mathbf{G}_i .

Network Representation (continued)

- Suppose that externalities take the following form

$$y_i = \alpha x_i + \gamma \mathbf{G}'_i \mathbf{x} + \mathbf{Z}'_i \boldsymbol{\beta} + \varepsilon_i, \quad (1)$$

where x_i is a measure of the human capital of individual i (e.g., schooling), and y_i is the outcome of interest (e.g., log earnings).

- Now it is straightforward to see that equation (1) enables a representation of the externalities and peer effects discussed over the last four lectures.
- For example:
 - Lucas (1988): \mathbf{G} 's all non-diagonal elements are equal to 1 (or the same number).
 - Other examples of externalities in the labor market or in schooling are more local, and necessitate different forms of \mathbf{G} .

Example

- As discussed in the first lecture, Acemoglu and Angrist derived and estimated an equation similar to (1) representing log wages (i.e., $y_i = \log \text{wages}$) with human capital externalities at the state level.
- How will individuals choose their schooling (their x_i) in this setting.
- First suppose that individuals just maximize log income minus costs of schooling. Then since log income is $\alpha x_i + \gamma \mathbf{G}'_i \mathbf{x} + \mathbf{Z}'_i \boldsymbol{\beta} + \varepsilon_i$, this maximization problem is

$$\max_{x_i} \alpha x_i + \gamma \mathbf{G}'_i \mathbf{x} + \mathbf{Z}'_i \boldsymbol{\beta} + \varepsilon_i - c(x_i).$$

- The solution to this problem will be related to the cost function and the return to schooling, captured by the parameter α :

$$\alpha = c'(x_i).$$

Example (continued)

- But suppose that instead of maximizing log income, individuals maximize actual income minus costs (just for illustration purposes).

$$\max_{x_i} e^{\alpha x_i + \gamma \mathbf{G}'_i \mathbf{x} + \mathbf{Z}'_i \boldsymbol{\beta} + \varepsilon_i} - c(x_i).$$

- In this case, the best-response equation of each individual—supposing that it is interior—would be of the form:

$$x_i = F(\mathbf{G}'_i \mathbf{x}, \dots),$$

thus leading to endogenous effects.

- Importantly, even though we started with a model without any endogenous effects, equilibrium behavior leads to endogenous effects.
- More generally, the resulting interactions will lead to game over networks.

Games over Networks

- Consider an economy consisting of a set of agents $N = \{1, 2, \dots, n\}$, with each agent's outcome equation given by the following quadratic:

$$y_i = (\alpha_i + \xi_i)x_i + \gamma_i \mathbf{G}'_i \mathbf{x} + \phi x_i \mathbf{G}'_i \mathbf{x} + \varepsilon_i, \quad (2)$$

where $x_i \geq 0$ is characteristics/choice of agent i , and ε_i and ξ_i are two random error terms (the role of each of them will become clear below). Think of ξ_i as a random variable.

- Note also that there are no *endogenous effects* in this equation.
- This is different from (1) in allowing:
 - Heterogeneous and random **own effects** (the coefficient of x_i)
 - Heterogeneous **spillovers** (the coefficient of $\mathbf{G}'_i \mathbf{x}$)
 - The possibility of **strategic effects** (strategic complements or substitutes through via the term $x_i \mathbf{G}'_i \mathbf{x}$).
- We discuss non-linearities below.

Games over Networks (continued)

- Agent i 's payoff is then obtained by introducing a quadratic cost:

$$U_i(x_i, \mathbf{x}_{-i}; \mathbf{G}) = y_i - \frac{\theta}{2}x_i^2.$$

- See Ballester, Calvo-Armengol, and Zenou *Econometrica*, 2006, or Bramoulle, Kranton, and D'Amours, *AER*, 2013.
- Applications (though with different functional forms):
 - private provision of public goods or jointly beneficial effort in networks (Bramoulle and Kranton *JET*, 2007, Allouch, mimeo 2013);
 - education decisions in social networks (Calvo-Armengol, Patacchini, and Zenou *Review of Economic Studies*, 2009)
 - oligopolistic competition (where the network represents substitution patterns);
 - innovation networks (Bramoulle, Kranton, and D'Amours, 2013);
 - crime networks (Calvo-Armengol, and Zenou, *International Economic Review*, 2004);
 - security investments against contagion (Acemoglu, Malekian, and Ozdaglar, mimeo 2013).

Best Responses

- Best responses are obtained straightforwardly as:

$$x_i = \max \left\{ \frac{\alpha_i}{\theta} + \frac{\phi}{\theta} \mathbf{G}'_i \mathbf{x} + \frac{\xi_i}{\theta}, 0 \right\}. \quad (3)$$

- This equation also clarifies the role of the ξ_i term in (2)—multiplicative random terms in the outcome equation translate into the error term in the best response equation.
- If $\phi > 0$, then this is a game of **strategic complements**—agents would like to take higher actions when their neighbors are doing so.
- If $\phi < 0$, then this is a game of **strategic substitutes**—agents would like to take lower actions when their neighbors are taking higher actions.

Nash Equilibria

- If all best responses were interior, the Nash equilibrium of this game could be written as a solution to the following matrix equation

$$\mathbf{x} = \tilde{\alpha} + \frac{\phi}{\theta} \mathbf{G}'_i \mathbf{x} + \tilde{\xi},$$

where $\tilde{\alpha}$ is a column vector with elements given by α_i/θ and $\tilde{\xi}$ is a column vector with elements given by ξ_i/θ .

- Assuming that the inverse $\left(\mathbf{I} - \frac{\phi}{\theta} \mathbf{G}'_i\right)^{-1}$ exists, the unique solution is

$$\mathbf{x} = \left(\mathbf{I} - \frac{\phi}{\theta} \mathbf{G}'_i\right)^{-1} (\tilde{\alpha} + \tilde{\xi}). \quad (4)$$

- Thus there is a unique **interior Nash equilibrium**—under the assumption that the inverse exists (or that ϕ/θ is not too large).
- There may be multiple non-interior equilibria, however.
 - E.g., with strategic substitutes, when an agent's neighbor takes a high action, she would like to take a low action, and vice versa.

Nash Equilibrium and Bonacich Centrality

- One popular measure of “centrality” of a node in the network is the Katz-Bonacich centrality or simply the Bonacich centrality measure (related to eigenvector centrality measures).
- In a network with adjacency matrix \mathbf{G} and a scalar a such that the matrix $(\mathbf{I} - a\mathbf{G})^{-1}$ is well-defined and nonnegative, the vector of Bonacich centralities of parameter a in this network is $(\mathbf{I} - a\mathbf{G})^{-1}\mathbf{e}$, where \mathbf{e} is the vector of 1's.
- Of course,

$$(\mathbf{I} - a\mathbf{G})^{-1} = \sum_{k=0}^{\infty} a^k \mathbf{G}^k,$$

where \mathbf{G}^k is the k th power of matrix \mathbf{G} (with $\mathbf{G}^0 = \mathbf{I}$).

Nash Equilibrium and Bonacich Centrality (continued)

- Intuitively, \mathbf{G}^k measures the number of “walks” of length k , with $g_{ij}^{[k]}$ as the number of walks of length k from i to j (think of the special case where this is a matrix of zeros and ones.)
- The parameter a is a decay factor that scales down the effect of longer walks.
 - In “generalized Bonacich” centrality measure a can be negative.
- Bonacich centrality of node i then counts the total number of walks that start from node i .
- This centrality measure (or its close cousins) emerge naturally in the equilibria of many linear, log-linear or quadratic models as (4) illustrates.
 - In particular, in this interior equilibrium, each agent’s action is equal to its “generalized” Bonacich centrality measure (the appropriate entry of $(\mathbf{I} - \frac{\phi}{\theta} \mathbf{G}'_i)^{-1}$) times its own effect, $\alpha_i/\theta + \xi_i/\theta$.

Nash Equilibria (continued)

- For the general characterization of Nash equilibria, let us call agents choosing strictly positive x_i *active agents*.
- Without loss of any generality, take the active agents to be those indexed $A = \{1, 2, \dots, a\}$, and partition \mathbf{G} as follows

$$\mathbf{G} = \begin{pmatrix} \mathbf{G}_A & \mathbf{G}_{A,N-A} \\ \mathbf{G}_{N-A,A} & \mathbf{G}_{N-A} \end{pmatrix}.$$

- Here \mathbf{G}_A is the matrix of the impact of active agents on active agents, and $\mathbf{G}_{N-A,A}$ of active agents on non-active agents.
- Then any Nash equilibrium is characterized by

$$\begin{aligned} \left(\mathbf{I} - \frac{\phi}{\theta} \mathbf{G}'_A\right) \mathbf{x}_A &= (\tilde{\alpha} + \tilde{\xi})_A, \\ \left(\mathbf{I} - \frac{\phi}{\theta} \mathbf{G}'_{N-A,A}\right) \mathbf{x}_A &\leq (\tilde{\alpha} + \tilde{\xi})_{N-A}. \end{aligned}$$

- I.e., given the actions of active agents, other active agents have an interior solution and non-active agents are happy at the corner.

Uniqueness

- Recall that a (best-response) *potential game* is a game that admits a potential function Γ , such that the derivative of this function with respect to the strategy/action of each player gives that player's best response.
- The game studied here is a potential game with potential function

$$\Gamma(\mathbf{x}, \mathbf{G}) = \mathbf{x}' (\tilde{\mathbf{a}} + \tilde{\boldsymbol{\zeta}}) - \frac{1}{2} \mathbf{x}' \left(\mathbf{I} - \frac{\phi}{\theta} \mathbf{G} \right) \mathbf{x}.$$

- This can be verified by checking that indeed $\frac{\partial U_i}{\partial x_i} = \frac{\partial \Gamma}{\partial x_i}$ for all i .

Uniqueness (continued)

- Sufficient condition for a unique equilibrium is a *strictly concave potential*, which will lead to a unique solution to the potential maximization problem, and thus to a unique Nash equilibrium.
- The above potential function is strictly concave if the matrix $\left(\mathbf{I} - \frac{\phi}{\theta} \mathbf{G}\right)$ is positive definite, which holds if its lowest eigenvalue is strictly positive.
- The eigenvalues of this matrix are simply given by 1 plus ϕ/θ times the eigenvalues of the matrix \mathbf{G} .
- Hence, a sufficient condition for uniqueness is $|\lambda_{min}(\mathbf{G})| < \theta/\phi$ (where $\lambda_{min}(\mathbf{G})$ is the smallest eigenvalue of \mathbf{G}).
- Thus, if $|\lambda_{min}(\mathbf{G})| < \theta/\phi$, then there is a unique Nash equilibrium.

Conceptual Problems with Endogenous Effects

- The best response equation, (3), clarifies that once we specify games over networks, *endogenous effects* are unavoidable.
 - These best responses link one agent's choice of x_i to another's choice.
 - Thus this is equivalent to endogenous effects now in x 's.
 - This is even though there were no endogenous effects in the original model.
- But does it make conceptual sense to say that my choice today depends on your choice today? How can something that has not been realized yet influence what you have chosen?
- Two answers to this:
 - This is what game theory predicts. But it's not the actual choice of yours that matters, but my anticipation of your choice.
 - Dynamics and stationary distributions...

Endogenous Effects and Stationary Distributions

- Consider the following dynamic variation:

$$y_{i,t} = (\alpha_i + \xi_i)x_{i,t} + \gamma_i \mathbf{G}'_i \mathbf{x}_t + \phi x_{i,t} \mathbf{G}'_i \mathbf{x}_t + \varepsilon_{i,t},$$

where the environment is not changing, but decisions are potentially time-varying.

- Suppose that agents best respond to the distribution of actions among other agents in the previous period.
- Then, the best-response equation is replaced by the following counterpart:

$$x_{i,t} = \max \left\{ \frac{\alpha_i}{\theta} + \frac{\phi}{\theta} \mathbf{G}'_i \mathbf{x}_{t-1} + \frac{\xi_i}{\theta}, 0 \right\}.$$

- This defines a *dynamical system*, and is easy to interpret—your decision from yesterday impacts my decision from today, so that it's only pre-determined variables that affect current choices.

Endogenous Effects and Stationary Distributions (continued)

- Moreover, under the sufficient condition for uniqueness, i.e., $|\lambda_{\min}(\mathbf{G})| < \theta/\phi$, but in fact more generally, this dynamical system converges to a *stationary distribution*, which will be given exactly by the best response equation, (3).
- Therefore, empirical equations that involve endogenous effects can alternatively be interpreted as representing the stationary distribution of a dynamic model.

Generalizations

- Games over networks can be specified with fairly general non-linear payoffs.
- A general treatment is in Allouch (2013).
- Most qualitative insights generalize to this non-linear setup; e.g., other centrality-type measures become important in this case.
- Uniqueness and stability condition can be generalized to

$$1 + \frac{1}{\lambda_{\min}(\mathbf{N}(\delta))} < \left(\frac{\partial x_i}{\partial \mathbf{G}'_i \mathbf{x}} \right)^{-1} < 1.$$

- Most importantly, estimation can be performed using non-linear methods using the structure of the game in a way that parallels the structural approach in the linear-quadratic case we discuss below.
- Similarly, incomplete information (especially about network structure) can also be introduced into this framework—see Galeotti, Goyal, Jackson, and Yariv (*Review of Economic Studies*, 2010).

Challenge I: Correlated Effects

- Consider the regressions

$$y_i = b_{\text{own}}x_i + b_{\text{spillover}}\bar{x}_i + \text{controls} + u_i^x, \quad (5)$$

$$y_i = b_{\text{own}}x_i + b_{\text{spillover}}\bar{y}_i + \text{controls} + u_i^y, \quad (6)$$

- where \bar{x}_i is the average of i 's neighbors, and \bar{y}_i is defined similarly.
- But unobserved errors are likely to be correlated between “neighbors”.
- Or in terms of (5) or (6), u_i^x and u_i^y are likely to be correlated across i .
- This is for two distinct but related reasons:
 - Suppose friendships are exogenously given. Two friends are still likely to be influenced by similar taste shocks, information and influences (the common shocks discussed above).
 - Suppose friendships are endogenously given. Then people choosing to be friends are likely to share similar observed and unobserved characteristics.

Challenge II: Endogenous Choices

- Somewhat less obvious, the fact that x 's are choices makes identification of certain parameters impossible—without taking a more “structural” approach.
- For this, take the best response equation of player i , (3) which, ignoring the corner solution, can be written as

$$x_i = \frac{\alpha_i}{\theta} + \frac{\phi}{\theta} \mathbf{G}'_i \mathbf{x} + \frac{\xi_i}{\theta}.$$

- Substitute this into the outcome equation, (2), which is

$$y_i = (\alpha_i + \xi_i)x_i + \gamma_i \mathbf{G}'_i \mathbf{x} + \phi x_i \mathbf{G}'_i \mathbf{x} + \varepsilon_i.$$

- We obtain

$$y_i = \theta x_i^2 + \gamma_i \mathbf{G}'_i \mathbf{x} + \varepsilon_i. \quad (7)$$

- Thus given endogenous choices, the crucial parameters of (2), in particular own effect (the average of α_i 's) and the strategic effect (ϕ), *cannot be identified*.

Challenge III: Endogenous Choices (continued)

- (Here, “identification” refers not to lack of identification of the regression coefficient (of say x_i or of y_i on the x choices of some neighbors), but to lack of information from an estimation approach on the “structural” or “causal” parameters).
- But all of these effects *can be* identified if (3) and (2) are estimated together:
- The estimation of (2) identifies the spillover effects—the average of γ_i 's—and the cost parameter θ .
- Knowing θ , the estimation of (3) identifies ϕ from the slope of the endogenous effects, and the average of the α_i 's is identified from the intercept.
- This underscores the importance of estimating the endogenous and contextual effects together when these relationships are derived from game-theoretic interactions (over networks).

Empirical Approaches

- Two approaches beyond the full random assignment (recall that with full random assignment and full compliance, there are no endogenous choices, so the issues discussed here are less relevant):
 - ① Exploit network structure.
 - ② “Network instruments”.
- Both approaches assume that the network structure is known and measured without error.

Estimation Problem

- Consider the two equations of interest—which might be jointly estimated.
- These do not include covariates, but α_i 's could be specified as functions of covariates and “excluded instruments”. In particular, suppose that

$$\alpha_i = h(\mathbf{z}_i' \boldsymbol{\beta}_x + \omega c_i),$$

where $h(\cdot)$ is a potentially non-linear function.

- Then

$$x_i = \frac{\phi}{\theta} \mathbf{G}_i' \mathbf{x} + \frac{1}{\theta} h(\mathbf{z}_i' \boldsymbol{\beta}_x + \omega c_i) + \tilde{\zeta}_i, \quad (8)$$

where $\tilde{\zeta}_i = \zeta_i / \theta$.

- The covariates \mathbf{z}_i can also be included in the outcome equation to obtain (for example, by specifying $\varepsilon_i = \mathbf{z}_i' \boldsymbol{\beta}_y + \tilde{\varepsilon}_i$).

$$y_i = \theta x_i^2 + \gamma_i \mathbf{G}_i' \mathbf{x} + \mathbf{z}_i' \boldsymbol{\beta}_y + \tilde{\varepsilon}_i. \quad (9)$$

Exploiting Network Structure

- The most well-known example of exploiting network structure is the creative paper by Bramoulle, Djebbari, and Fortin (*Journal of Econometrics*, 2009).
- Consider three agents i, j and k , and let us use the notation $k \in N(j)$ to denote that k is linked to (is a neighbor of) j .
- Suppose that $k \in N(j)$, $j \in N(i)$, and $k \notin N(i)$ —i.e., k is j 's friend/neighbor and j is i 's friend/neighbor, but k is not links to i .
- Then, in terms of estimating (8) for of the impact of x_j on x_i , we can use covariates of k , \mathbf{z}_k , as instruments.

Exploiting Network Structure (continued)

- But this identification strategy works only if error terms in the best response and the outcome equations, (8) and (9)— $\tilde{\xi}_i$ and $\tilde{\epsilon}_i$ —orthogonal across non-neighbor agents.
 - Bramoullé et al. show how one might deal with some instances of a *a priori known* correlated effects.
- If k and i have correlated error terms that are also correlated with their characteristics (their x 's), then k 's covariates cannot be an instrument for estimating j 's endogenous effect on i .
- But such correlation is likely to be endemic:
 - Geographic or social proximity between k and i likely to be high because they share friends.
 - Unlikely that k and j are correlated, j and i are correlated, but k and i are uncorrelated.
- Additional problem: if the network is measured with error, then neighbors k and i may appear not to be neighbors, creating a violation of the exclusion restrictions.

Network Instruments

- Suppose that there is a variable—unrelated to the network— c_i orthogonal to $\tilde{\zeta}_i$ and $\tilde{\varepsilon}_i$ that can be used as an instrument for x_i absent any externalities, peer effects or network interactions.
- Then this is a candidate to be a variable that is orthogonal to $\tilde{\zeta}_k$ and $\tilde{\varepsilon}_k$ for all $k \neq i$.
- In other words, if we have

$$\text{cov}(c_i, \tilde{\zeta}_i) = \text{cov}(c_i, \tilde{\varepsilon}_i) = 0,$$

then it is also plausible that (for any integer p)

$$\text{cov}(\mathbf{G}'_i \mathbf{c}, \tilde{\zeta}_i) = \text{cov}((\mathbf{G}_i^p)' \mathbf{c}, \tilde{\zeta}_i) = \text{cov}(\mathbf{G}'_i \mathbf{c}, \tilde{\varepsilon}_i) = \text{cov}((\mathbf{G}_i^p)' \mathbf{c}, \tilde{\varepsilon}_i) = 0.$$

Network Instruments (continued)

- But c_i should ideally satisfy an additional condition: *lack of correlation over the network*, i.e.,

$$\text{cov}(\mathbf{c}, (\mathbf{G}_i^p)'\mathbf{c}) \approx 0.$$

- Why?
- Because, otherwise, the correlated unobserved effects $\tilde{\zeta}_i$ and $\tilde{\varepsilon}_i$ could project onto \mathbf{c} .

Estimating the Network

- Does it matter if the network is not known?
- Yes and no.
- If there is no information on the network, then instead of a single parameter ϕ or a well-defined local average of γ_i 's, we would need to estimate $n(n-1)$ parameters, which is not feasible.
- But if the network is known up to some parameter δ , that parameter (or parameter vector) can also be consistently estimated.

Estimation

- Finally, estimation could be performed by instrumental-variables separately on (8) and (9).
- But additional efficiency can be gained by estimating these equations jointly by GMM or other methods.
- This is particularly true if there is a parameter of the network, δ , also to be estimated.

An Application

- The only application is from a non-peer effects setting, from Acemoglu, Garcia-Jimeno, and Robinson (2013).
- x_i = state capacity (presence of state agencies and employees) at the municipality level in Colombia.
- y_i = prosperity, poverty etc. at the municipality level.
- $\mathbf{G}(\delta)$ = the municipality network given by distances and variance of elevation between municipalities (with the parameter δ corresponding to the weighing of distances and elevation).
- c_i = historical variables on where the Spanish set up the colonial state and also on the road network they constructed built on income roads, which later disappeared.
- \mathbf{z}_i = various controls, including population.

Context

- General agreement that the weakness of the state and lack of economic integration has been a major problem in Colombian history and economic development.
- Country split by the Andes creating relatively isolated subregions.
- Colonial state concentrated in a few places and absent from much of the rest of the country.
- In the 19th century, number of public employees relative to population about 1/10 of contemporary US level.
- Rufino Gutierrez in 1912:

“...in most municipalities there was no city council, mayor, district judge, tax collector... even less for road-building boards, nor whom to count on for the collection and distribution of rents, nor who may dare collect the property tax or any other contribution to the politically connected...”

- The same seems to be true today even if to a lesser extent.

Model Setup

- Identical to the game over networks presented above with the adjacency matrix given by

$$g_{ij} = \begin{cases} 0 & \text{if } j \notin N(i) \\ f_{ij} & \text{if } j \in N(i) \end{cases}$$

where

$$f_{ij} = \frac{1}{1 + \delta_1 d_{ij}(1 + \delta_2 e_{ij})}.$$

- $N(i)$ is the set of neighbors of i , d_{ij} is geodesic distance between i and j , e_{ij} is variability in altitude along the geodesic.
- The rest of the model is the same and our identification strategy will be the same as the one described above.

Instruments

- *Colonial state presence*, measured either by the number of colonial state agencies or employees.
 - Highly concentrated around key cities and resources, including military aims in strategic places.
 - In gold mining regions, colonial state presence related to taxation.
 - In high native population regions, related to control of the population, legal adjudication, etc.
 - Gold mining, native populations and those military aims are no longer relevant. So the direct effect of colonial state presence is by creating the infrastructure for current state presence.
- *Royal roads* were one of the few investments in infrastructure (building upon pre-colonial roads).
 - The presence of royal roads is a good indicator of where the colonial state was interested in reaching out, and controlling territory.
 - But most of these royal roads were subsequently abandoned as transportation infrastructure.

The Correlation Matrix

- These measures are not geographically correlated (reflecting the specific Spanish colonial strategy).

Table 2. Within-department Spatial Correlation of Historical State Presence Variables

	1	2	3	4	5	6
1. Own distance to royal roads	1.000					
2. Neighbors' average distance to royal roads	0.283	1.000				
3. Own colonial officials	-0.095	-0.072	1.000			
4. Neighbors' average colonial officials	-0.146	0.039	-0.061	1.000		
5. Own colonial state agencies	-0.135	-0.039	0.545	-0.006	1.000	
6. Neighbors' average colonial state agencies	-0.208	0.250	-0.053	0.490	0.022	1.000

Correlations reported are the average across-departments of the correlations for each department.

Estimates of the Best Response Equation

- Fix δ the specific value, here $(1, 1)$, and estimate the two equations separately by linear IV or by GMM (also estimating δ).
- The sufficient condition for uniqueness are always satisfied.
- Robust evidence for **strategic complementarities**, i.e., $\phi > 0$.

Estimates of the Best Response Equation

Table 3. Contemporary State Equilibrium Best Response

State capacity measured as log of:	Number of municipality employees			
Panel I	(5)	(6)	(7)	(8)
	OLS	IV	IV	Sys. GMM
dx_i/dx_j	0.021 (0.003)	0.022 (0.004)	0.022 (0.004)	0.016 (0.003)
$dx_i/d\text{colonial state officials}_i$	0.129 (0.043)	0.130 (0.043)	0.105 (0.047)	0.087 (0.069)
$ds_i/d\text{colonial state agencies}_i$	0.017 (0.058)	0.017 (0.059)	-0.002 (0.060)	0.085 (0.085)
$dx_i/d\text{distance to royal road}_i$	-0.035 (0.034)	-0.035 (0.035)	-0.038 (0.036)	-0.036 (0.044)
Panel II		First Stage		
First-stage R-squared:		0.681	0.658	
F-test for excluded instruments:		19.55	171.0	
F-test p-value		0.000	0.000	
Overidentification test:	Test statistic	4.399	5.775	
Chi-squared(2)	P-value	0.494	0.449	
Log population	Control	Control	Instrum	Instrum
Observations	1017	1017	1017	1003

Estimates of the Outcome Equation

Table 4B. Prosperity and Public Goods Structural Equation

State capacity measured as: log of number of municipality employees

Dependent variable	% Not in poverty			
Panel I	(9)	(10)	(11)	(12)
	OLS	IV	IV	Sys. GMM
	Prosperity equation			
dy_i/dx_i	0.233 (0.021)	0.305 (0.119)	0.275 (0.111)	0.210 (0.023)
dy_i/dx_j	0.019 (0.003)	0.013 (0.005)	0.014 (0.005)	0.016 (0.002)
	First stage for x_i^2			
F-test for excluded instruments:		13.28	27.42	
F-test p-value		0.000	0.000	
First-stage R-squared		0.570	0.575	
	First stage for $G_i(\delta)'x$			
F-test for excluded instruments:		344.4	457.4	
F-test p-value		0.000	0.000	
First-stage R-squared		0.759	0.758	
Log population	Control	Control	Instrum	Instrum

Quantitative Magnitudes

- One advantage of estimating the structural model is the ability to perform counterfactual exercises, taking equilibrium into account.
- The larger effects in Panel Ib because of “network effects” (and strategic complementarities).

**Table 5. Experiment:
Implications of Moving All Municipalities below Median State Capacity to Median**

Panel Ia	Linear model			
	Municipality Employees:		% not in poverty	
Partial equilibrium change in:	From	To	From	To
Change in median:	20	20	57.1	60.0
Fraction due to own effect:				57.1%
Fraction due to spillovers:				43.0%
Panel Ib				
General equilibrium change in:	Municipality Employees:		% not in poverty	
	From	To	From	To
Change in median:	20	42.3	57.1	68.3
Fraction due to direct effect:				25.5%
Fraction due to network effects:				74.5%

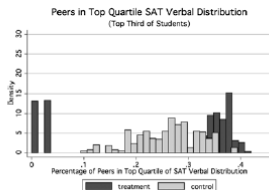
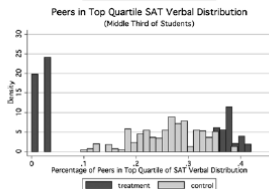
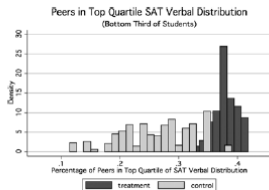
Endogenous Networks

- A large literature studies the endogenous formation of (social) networks—e.g., Jackson and Wolinsky (*JET*, 1996), Bala and Goyal (*Econometrica*, 2000).
- Endogeneity of networks makes externalities and peer effects more interesting but also more complicated conceptually and more difficult to estimate.
- Application: Carrell, Sacerdote, and West (*Econometrica*, 2013).
- Estimate peer effects across cadets within squadrons using random assignment from the U.S. Air Force Academy.
- These peer effects were non-linear:
 - Low (baseline) ability students appeared to benefit significantly from being in the same squadron as high-ability students with limited negative effect on high-ability students from such mix.
- This suggests that optimally manipulating the composition of squadrons can lead to significant gains.

Endogenous Networks: Intervention Gone Wrong

- The authors convinced the U.S. Air Force Academy to allow such manipulation, and constructed “optimally designed” squadrons—in which the exposure of low-ability cadets to high-ability ones was maximized by creating “*bimodal*” squadrons.
- However, instead of the hypothesized gains, there were losses among low-ability cadets. Why?
- The authors hypothesize, and provides some evidence in favor of, the following story:
 - The real peer groups—the friendship networks—probably changed as a result of the intervention: low-ability and high-ability cadets may have stopped working and being friends together in the bimodal squadrons.
 - As a result, the peer effects from high-ability to low-ability cadets weakened or disappeared, leading to negative results.
- A cautionary tale on the endogeneity of social networks with respect to interventions.

Endogenous Networks: Bimodal Treatment



Endogenous Networks: Prediction Vs. Realization

TABLE IV
PREDICTED TREATMENT EFFECT^a

Variables	(1) All Students	(2) Bottom \widehat{GFR}	(3) Middle \widehat{GFR}	(4) Top \widehat{GFR}
Student in Treatment Group	2.787 (0.026)	2.390 (0.027)	2.783 (0.027)	3.198 (0.027)
Student in Control Group	2.772 (0.026)	2.336 (0.027)	2.767 (0.027)	3.195 (0.026)
Predicted Treatment Effect	0.015 (0.037)	0.053 ^b (0.037)	0.016 (0.037)	0.003 (0.037)
Observations	2653	881	884	888

TABLE VI
OBSERVED TREATMENT EFFECTS^a

Variables	(1) All Students	(2) Low \widehat{GFR}	(3) Middle \widehat{GFR}	(4) High \widehat{GFR}
Student in Treatment Group	0.001 (0.022)	-0.061 ^c (0.031)	0.082 ^b (0.039)	-0.012 (0.036)
Observations	4834	1571	1626	1637
R^2	0.357	0.136	0.067	0.151

Endogenous Networks: A Possible Explanation?

TABLE VIII
LOW PREDICTED GPA STUDENTS: TREATMENT EFFECTS ON STUDY PARTNER AND
FRIEND CHOICES*

Treatment Effect On ...	(1)	(2)	(3)	(4)	(5)	(6)
	Study Partners			Friends		
	Actual Peer Choices (sd)	If Peers Chosen Randomly (sd)	Actual Minus Random $P(A < R)$	Actual Peer Choices (sd)	If Peers Chosen Randomly (sd)	Actual Minus Random $P(A < R)$
Fraction Low \widehat{GPA}	0.171 ^b (0.061)	0.125 ^b (0.020)	0.046 ^c 0.011	0.201 ^b (0.050)	0.135 ^b (0.019)	0.071 ^b 0.000
Fraction Middle \widehat{GPA}	-0.214 ^b (0.046)	-0.105 ^b (0.015)	-0.109 ^b 1.000	-0.105 ^b (0.038)	-0.114 ^b (0.014)	0.008 0.272
Fraction High \widehat{GPA}	0.042 (0.060)	-0.020 (0.016)	0.062 ^b 0.000	-0.095 ^c (0.048)	-0.022 (0.015)	-0.074 ^b 1.000
Fraction High SAT-V	0.064 (0.052)	0.103 ^b (0.015)	-0.039 ^c 0.994	0.004 (0.046)	0.112 ^b (0.014)	-0.107 ^b 1.000
Fraction Low $\widehat{GPA} > 0.50$	0.269 ^b (0.097)	0.170 ^b (0.042)	0.100 ^c 0.008	0.236 ^c (0.092)	0.170 ^b (0.041)	0.065 0.057
Observations	494	10,000	10,000	543	10,000	10,000