Networks: Games over Networks and Peer Effects

Daron Acemoglu

MIT

July 22, 2014.

Introduction

- The potential importance of various different types of externalities (e.g., human capital externalities) and peer effects are well recognized.
- A growing empirical literature investigates the extent of such effects in different contexts.
- Networks give a powerful language for the systematic theoretical and empirical investigation of such externalities.
- But even more importantly, the endogeneity (and the associated identification) challenges can be best understood by considering economic decisions in a network setting.
- Finally, network-based approaches can also clarify and help the empirical challenges in this area.

Plan

- Networks and externalities.
- Games over networks and "peer effects".
- "Identification problems".
- Empirical approaches.
- Application.
- What has been left out of this lecture: endogenous network formation and empirical implications.
- Conclusion.

Networks and Externalities

- Consider the problem of human capital externalities.
- But who creates externalities on whom?
  1. Lucas (*Journal of Monetary Economics*, 1988): human capital externalities from each worker to everybody else.
  2. Marshall (*Principles of Economics,* 1920): human capital externalities from workers to others in the same location and industry.
  3. Jane Jacobs (*The Economy of the City,* 1970): human capital externalities within cities, and in her more specific version, between managers in different industries within the same city.
  4. Rauch (*Journal of Urban Economics,* 1993): human capital externalities within cities across workers. Acemoglu and Angrist (*NBER Macroeconomics Annual,* 2002), Moretti (*Journal of Econometrics,* 2004): human capital externalities within states across workers.
  5. etc.

- The language of networks useful for representing and thinking about these issues.

Network Representation

- Consider the following reduced-form model in an economy consisting of n individuals/workers.
- A (nonnegative) matrix **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 **G** are nonnegative numbers (normalized to be less than 1), and **G** is not necessarily symmetric.
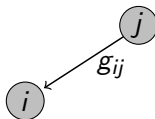
- We also adopt the convention that *the diagonal elements of G are zero*, and we denote the entries by g_{ij} and the ith column of **G** by $\mathbf{G}_i$.

Network Representation (continued)

- Suppose that externalities take the following form

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

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 the representation of the aforementioned externalities in terms of the simple equation.

- For example:
  - Lucas (1988): $\mathbf{G}$'s all non-diagonal elements are equal to 1 (or the same number).
  - The other examples implicitly specify a network structure (e.g., workers are neighbors to others in the same industry and location in Marshall's model) and this can be captured, of course, by the appropriate $\mathbf{G}$.

Contextual and Endogenous Effects

- In Manski's (*Review of Economic Studies*, 1993) language equation (1) specifies **contextual effects**.
  - Put more simply, externalities are from **pre-determined** characteristics.
- The alternative is the so-called **endogenous effects**, where the relationship would take the form

$$y_i = \alpha x_i + \tau \mathbf{G}_i' \mathbf{y} + \mathbf{Z}_i' \boldsymbol{\beta} + \varepsilon_i, \tag{2}$$

and of course the two could be combined.

- One view is that equation (2) is more challenging empirically and less compelling conceptually than (1)—e.g., Angrist (2004); more on this below.
- Of course, one could solve out for y_i starting from (2) to arrive to a version of (1).
- But more importantly, unless x_i is not just pre-determined but randomly assigned, its determination will lead to a form similar to (2)—except in x's rather than in y's as we will see.

Example

- In Acemoglu and Angrist (2002), we derived and estimated an equation similar to (1) representing log wages with human capital externalities at the state level.

- We used an instrumental-variables (IV) strategy for dealing with endogeneity of both own and state-level average schooling.

- The special assumptions—e.g., individuals maximize log income minus costs—made there ensured that each individual chose her own schooling independent of schooling level of all others in the economy.

Example (continued)

- But this need not the the case.
- Suppose, e.g., each individual maximizes income—rather than log income—minus cost of schooling.

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

- Or there are interactions between the rate of return on human capital and human capital of others.
- In that case, the **best-response** equation of each individual—supposing that it is interior—would be of the form:

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

thus leading to **endogenous effects**.

- More generally, the resulting interactions will lead to a **game over networks**.

Games over Networks

- Consider an economy consisting of a set of agents $N = \{1, 2, ..., n\}$.
- Each agent's **outcome equation** is assumed to take the following quadratic form:

$$y_i = (\alpha_i + \xi_i)x_i + \gamma_i \mathbf{G}'_i \mathbf{x} + \phi x_i \mathbf{G}'_i \mathbf{x} + \varepsilon_i, \tag{3}$$

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). For now, note that α_i is known, whereas ξ_i is a random variable.

- 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.

A Note on Estimation

- In theory, all three of the relevant effects—local averages of own effects, spillovers and strategic effects—could be obtained by non-linearly estimating (3).
- But we will see that this is in general not feasible unless *all* x_i's are randomly assigned.
- Having an exogenous source of variation is typically not sufficient because of the endogeneity of the x_i's.
  - This is distinct from the challenges of estimating peer effects highlighted in Acemoglu and Angrist (2002), Boozer and Cacciola (mimeo, 2001) and Angrist (*Labour Economics,* 2014).

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:

$$x_i = \max\left\{ \frac{\alpha_i}{\theta} + \frac{\phi}{\theta}\mathbf{G}'_i\,\mathbf{x} + \frac{\xi_i}{\theta}, 0 \right\}. \tag{4}$$

- This equation also clarifies the role of the ξ_i term in (3)—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

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

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

$$\mathbf{x} = \left(\mathbf{I} - \frac{\phi}{\theta}\mathbf{G}_i'\right)^{-1}\left(\tilde{\boldsymbol{\alpha}} + \tilde{\boldsymbol{\xi}}\right). \qquad (5)$$

- 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 **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 **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 kth power of matrix **G** (with $\mathbf{G}^0 = I$).

Nash Equilibrium and Bonacich Centrality (continued)

- Intuitively, $\mathbf{G}^k$ measures the number of "walks" of length k.

    - $g_{ij}^{[k]}$ is the number of walks of length k from i to j.

- 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 (5) illustrates.

    - In particular, in this interior equilibrium, each agent's action is equal to its "generalized" Bonacich centrality measure (the appropriate entry of $\left(\mathbf{I} - \frac{\phi}{\theta}\mathbf{G}_i'\right)^{-1}$) 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 $A = \{1, 2, ..., a\}$, and partition **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

$$(\mathbf{I} - \frac{\phi}{\theta}\mathbf{G}_A')\mathbf{x}_A = (\tilde{\boldsymbol{\alpha}} + \tilde{\boldsymbol{\xi}})_A,$$

$$(\mathbf{I} - \frac{\phi}{\theta}\mathbf{G}_{N-A,A}')\mathbf{x}_A \leq (\tilde{\boldsymbol{\alpha}} + \tilde{\boldsymbol{\xi}})_{N-A}.$$

- 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 admit 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

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

- 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, (4), 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.

- 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, (4).

- Therefore, empirical equations that involve endogenous effects can alternatively be interpreted as representing the stationary distribution of a dynamic model.

Do the Simplifying Assumptions Matter?

- Yes and no.
- 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).

Identification Challenges

- The difficulty of identifying both contextual and endogenous effects are well recognized since Deaton (mimeo 1990) and Manski (1993).
- But the networks perspective clarifies the issues.
- Throughout, "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.

Challenge I: Mechanical Biases

- A regression of the form

$$y_i = b_{\text{own}} x_i + b_{\text{spillover}} \bar{x}_i + \text{ controls} + u_i^x, \tag{6}$$

  where $\bar{x}_i$ is the average of i's neighbors, will not always leads to interpretable estimates of $b_{\text{spillover}}$.

- Acemoglu and Angrist (2002): in the case where neighborhoods are "partitioned"— i.e., neighbors are other units in the same area—even if there are no spillovers, $b_{\text{spillover}}$ will be estimated to be positive *provided that* OLS estimates of own effects differ from IV estimates using group dummies (e.g., because of different local averages or because of measurement error).

- The problem is even worse when the regression takes the form

$$y_i = b_{\text{own}} x_i + b_{\text{spillover}} \bar{y}_i + \text{ controls} + u_i^y. \tag{7}$$

Challenge II: Correlated Effects

- Unobserved errors are likely to be correlated between "neighbors".
- Or in terms of (6) or (7), u_i^x and u_i^y are likely to be correlated across i.
- This is for two distinct but related reasons:

   1. Suppose friendships are exogenously given. Two friends are still likely to be influenced by similar taste shocks, information and influences (because they are spending time together or because roommates are affected by the same disturbances, a problem even for papers such as Sacerdote (*QJE*, 2001) attempting to estimate endogenous effects based on random assignment).
   2. Suppose friendships are endogenously given. Then people choosing to be friends are likely to share similar observed and unobserved characteristics.

- Empirical approaches outlined below will attempt to deal with both mechanical biases and correlated effects.

Challenge III: 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, (4) 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, (3), 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

$$y_i = \theta x_i^2 + \gamma_i\mathbf{G}'_i\,\mathbf{x} + \varepsilon_i. \tag{8}$$

- Thus given endogenous choices, the crucial parameters of (3), in particular own effect (the average of α_i's) and the strategic effect (ϕ), *cannot be identified*.

Challenge III: Endogenous Choices (continued)

- But all of these effects *can be* identified if (4) and (3) are estimated together:

- The estimation of (3) identifies the spillover effects—the average of γ_i's—and the cost parameter θ.

- Knowing θ, the estimation of (4) 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:
  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

$$x_i = \frac{\phi}{\theta} \mathbf{G}_i' \mathbf{x} + \frac{1}{\theta} h\left(\mathbf{z}_i' \boldsymbol{\beta}_x + \omega \mathbf{c}_i\right) + \tilde{\xi}_i, \qquad (9)$$

where $\tilde{\xi}_i = \xi_i / \theta$.

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

$$y_i = \theta x_i^2 + \gamma_i \mathbf{G}_i' \mathbf{x} + \mathbf{z}_i' \boldsymbol{\beta}_y + \tilde{\varepsilon}_i. \qquad (10)$$

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 (9) 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, (9) and (10)—$\tilde{\tilde{\zeta}}_i$ and $\tilde{\varepsilon}_i$—orthogonal across non-neighbor agents.
    - Bramoulle et al. show how one might deal with some instances of *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

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

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

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

Network Instruments (continued)

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

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

- Why?
- Because, otherwise, the correlated unobserved effects $\tilde{\tilde{\varsigma}}_i$ 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 (9) and (10).

- 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

- 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	0.022	0.022	0.016
	(0.003)	(0.004)	(0.004)	(0.003)
dx_i/dcolonial state officials_i	0.129	0.130	0.105	0.087
	(0.043)	(0.043)	(0.047)	(0.069)
ds_i/dcolonial state agencies_i	0.017	0.017	-0.002	0.085
	(0.058)	(0.059)	(0.060)	(0.085)
dx_i/ddistance to royal road_i	-0.035	-0.035	-0.038	-0.036
	(0.034)	(0.035)	(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	0.305	0.275	0.210
	(0.021)	(0.119)	(0.111)	(0.023)
dy_i/dx_j	0.019	0.013	0.014	0.016
	(0.003)	(0.005)	(0.005)	(0.002)
Panel II		**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
Observations	1017	1017	1017	1003

Quantitative Magnitudes

- One advantage of estimating the structural model is the ability to perform counterfactual exercises, taking equilibrium into account. This makes an important difference.

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

Panel Ia	Linear model			
Partial equilibrium change in:	Municipality Employees:		% not in poverty	
	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%	

From Linear to Non-Linear Models

- One can also incorporate nonlinear interactions, which become necessary if we wish to take account of the contribution of national state bureaucracy at the local level.

- Then the model can be estimated using maximum likelihood or simulated method of moments.

- The basic conclusions seem to be robust.

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 has high-ability students with limitednegative effect on high-ability students from such mixin.
- This suggests that optimally manipulating the composition of squadrons can lead to significant gains.

Endogenous Networks (continued)

- 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.

Conclusion

- Networks as a language and as a systematic method of building theoretical and empirical analyses of externalities, peer effects and strategic interactions.
- The presence of network interactions underscores with challenging nature of estimation of externalities, peer effects, spillovers etc.
- But the study of games over networks and network interactions might also provide promising directions for overcoming some of these quintessential conceptual and empirical challenges.