UNIVERSITY OF CALIFORNIA

Santa Barbara

Essays on Social Influences in Economic Decision Making

A Dissertation submitted in partial satisfaction
of the requirements for the degree of

Doctor of Philosophy

in

Economics

by

Daniel Wecker Argyle

Committee in Charge:

Professor Richard Startz, Chair

Professor Douglas G. Steigerwald

Professor Kelly Bedard

June 2014

The Dissertation of
Daniel Wecker Argyle is approved:

_____

Professor Douglas G. Steigerwald

_____

Professor Kelly Bedard

_____

Professor Richard Startz, Committee Chairperson

June 2014

Essays on Social Influences in Economic Decision Making

Copyright © 2014

by

Daniel Wecker Argyle

To Lisa, I wouldn't have survived it without you,

and Lydia and Isaac, for making it fun.

Acknowledgements

Finishing this work would not have been possible without the contributions of many, especially those on my committee. I thank my advisor Dick Startz for providing sound counsel for research, career, and life, and, most importantly, believing in me; Doug Steigerwald for extending new opportunities in teaching and research, as well as thoughtfully providing feedback and a smile; and Kelly Bedard for hours of conversation traveling and filling in adeptly and willingly. Although he is not able to sign this dissertation, I want to highlight to the influence of Phillip Babcock. His knowledge, kindness, and encouragement started me down this path. I will remember him fondly.

My time at UCSB has been wonderful, with far too many people contributing to the experience to identify them all individually. My fellow graduate students, especially those in my cohort, provided much needed support and friendship. I am especially grateful for those in the Econometrics Research Group and the Education Research Group. They provided detailed and useful feedback on all of the research that is contained in this document. I especially thank Stefanie Fischer for her enthusiasm and patience as we worked on our research about four-day school weeks.

Family is vitally important to me and they have helped in innumerable ways through my time in graduate school. I am grateful for them all. I especially thank my mother, Julie Argyle; she suggested I study economics in the first place. She also, along with my mother-in-law Deb Pfost, provided necessary support in the final days of finishing this work. Most important of all, I am grateful for my wife Lisa. Without her hours of conversation and editing, I would not be able to get any research done at all. She is a brilliant academic and an even better mother to our two children. Lastly, I thank Lydia and Isaac; they make my life happy.

Curriculum Vitæ
Daniel Wecker Argyle

Education

June 2014	Ph.D. Economics, University of California, Santa Barbara

- Fields: Econometrics, Labor Economics
- Committee: Dick Startz, Doug Steigerwald, Kelly Bedard

Dec. 2010	M.A. Economics, University of California, Santa Barbara
Aug. 2009	B.A. Economics, Brigham Young University, Provo, Utah

- University Honors and Mathematics Minor
- Honors Thesis: *Student Achievement in Charter Schools, A Cross-State Analysis*

Grants, Awards, and Fellowships

2013	Broom Center Graduate Student Travel Grant
2011	UCSB Political Science Manzer-Wesson Award for Best Graduate Student Paper (with Lisa Argyle)
2009-2010	Andron Graduate Fellowship - University of California, Santa Barbara
2008-2009	President, BYU Economics Student Association
2008	First Place, Mary Lou Fulton Mentored Research Poster Conference
2007	BYU Office of Research and Creative Activities Grant
2003	BYU Heritage Scholarship

Teaching Experience

Primary Instructor: Undergraduate Econometrics, Principles of Economics, Statistics for Economists

Teaching Assistant: Graduate Econometrics, Undergraduate Econometrics, Intermediate Microeconomics

Presentations University of Redlands (2014), U.S. Department of Justice (2014), Australian National University (2014), California Econometrics Conference (2013), All-California Labor Conference Poster Session (2013), LSU Conference on Networks (2013), Utah Conference for Undergraduate Research (2009)

Abstract

Essays on Social Influences in Economic Decision Making

Daniel Wecker Argyle

Understanding social influences is vitally important to understanding human behavior; unfortunately, observing and measuring these effects is extremely difficult. This dissertation focuses on applying and developing techniques to measure social influences on individual decision making. The first chapter develops two techniques for estimating social influences using friendship networks, even when individuals differ on unobservable attributes. The second chapter uses social networks sampled from villages in India to demonstrate strong peer influences on an individual's decision to take a microfinance loan. The final chapter examines changes in juvenile criminal behavior in response to their schools changing to a four-day school week. An abstract for each chapter is provided below.

Chapter 1 Abstract: Existing methods for identifying peer effects in social networks require the assumption that there are no unobserved factors that determine both network links and the outcome of interest. Since this assumption is likely to be violated in most networks, I provide two distinct methods for estimating peer effects in the presence of individual heterogeneity. First, I show that if repeated observations on individuals are available, accounting for time-invariant individual attributes can restore

identification. Second, I suggest a semi-parametric Bayesian instrumental variables technique that can be used to estimate peer effects models when panel data is not available or if fixed effects assumptions are undesirable. I demonstrate the properties of both strategies using simulated network data and provide evidence that both are able to estimate peer effects, even when networks exhibit sorting on observable and unobservable attributes. Additionally, I apply both techniques to estimate peer effects in the United States Congress using cosponsorship networks and voting outcomes.

Chapter 2 Abstract: I use friendship networks from villages in rural India to provide a careful examination of peer effects in decisions to take up microfinance. Using social networks resolves the traditional identification problems associated with peer effect estimation and allows each individual to have their own peer group. Since the networks in question are sampled and some friendship information is missing, I use an analytic correction to ensure unbiased estimates as well as a technique known as graphical reconstruction to provide bounds on estimators accounting for possible configurations of the unobserved links in the data. In addition to including a variety of individual covariates, I control for observable characteristics of an individual's peers, examine different ways of specifying network links, and account for the probability of being informed about the program. Across these specifications I find positive and significant peer effects; an additional friend participating in microfinance increases the probability that an individual participates in microfinance by 6%-10%.

Chapter 3 Abstract: There has been a recent trend to implement four-day school weeks, especially in rural areas. We show that changing to a four-day school week influences youth criminal activity in jurisdictions where these policies take effect, especially impacting property crime. Focusing on Colorado, where four-day school weeks are especially prevalent, we show that switching all students in a county to a four-day week leads to 0.75 additional juvenile arrests for larceny per 1,000 residents. Effects on other types of crime are inconclusive; estimated signs on drug and violent crimes are negative, although neither attains statistical significance.

<div style="text-align: right;">

Professor Richard Startz

Dissertation Committee Chair

</div>

Contents

---

[1] **Joint Work with Stefanie Fischer**

List of Figures

List of Tables

Chapter 1

Individual Heterogeneity and Peer Effects in Networks

1.1   Introduction

Peer effects emerge in many settings where a social group may influence an individual's outcome. Traditionally, economic research exploits some sort of exogenous variation in social group formation to examine effects as diverse as academic achievement (Sacerdote, 2001), health habits (Carrell et al., 2011), and income (Lerner and Malmendier, 2011). Recent work in econometrics has focused more on how peer effects can be identified even when social groups vary at the individual level and are selected by the individual. Bramoullé et al. (2009) establish conditions for the identification of peer effects from a social network.[1] In so doing, they require a strict exogeneity assumption that is violated if unobserved attributes affect both the network formation process and the outcome of interest. Since this is unlikely to be true in most networks, I generalize the exogeneity assumption to allow for unobserved time-invariant individual level attributes. If repeated observations on individuals are available, I show that identification of peer effects is restored under the more general assumption and that the parameters of the model can be consistently estimated using and instrumental variables procedure. In addition, I propose a semi-parametric estimation strategy that is robust to individual heterogeneity and may be appropriate in cases where panel assumptions are undesirable or panel data is not available. The semi-parametric technique has the advantage of flexibly accounting for individual heterogeneity and does not require the same assumptions that the panel

---

[1] De Giorgi et al. (2010) established similar results nearly simultaneously.

methods do. Both panel and semi-parametric estimation have different strengths and each may be appropriate depending on the data that is available; I demonstrate their strengths and weaknesses in simulated networks and in an application estimating peer effects in Congressional voting.

Goldsmith-Pinkham and Imbens (2013) also set out to account for individual unobservable characteristics that affect both the network formation process and the outcome using the Bramoullé et al. (2009) framework. They posit a specific network formation process that depends on a binary individual-level latent attribute that is also allowed to affect the outcome.[2] This approach has the advantage of only requiring a single network; however, this comes at the expense of nontrivial additional assumptions. Most notably, this procedure requires that the posited network formation process correctly specifies the true formation process. In contrast, the approach taken in this paper requires no specific assumptions about the structure of the network formation process, stipulating only that the network must be formed exogenously from the data generating process for the outcome.

There are two specific ways that networks are assumed to be exogenous. I first examine a scenario where the network is fixed and multiple outcomes are observed; in this case the structure of the network can be taken as given in the subsequent data

---

[2]Goldsmith-Pinkham and Imbens (2013) do consider a situation where they have two observations of a networks observed at distinct points in time; however, their primary concern is using the networks to estimate the specific network formation process suggested in their work. In addition, they also explore several interesting extensions of the model, most notably including spatial auto-regressive and moving-average models.

generating process for the outcomes. Second, I generalize peer effects estimation to a situation where the network is allowed to vary over time; however, the networks must be independent from one another conditional on modeled individual covariates. In the empirical application I examine voting behavior of legislators in a Congressional session where friendship links are determined by cosponsorship decisions. In this context, the first scenario corresponds to a case where cosponsorship decisions of legislators is determined separately from their voting decisions. To increase the plausibility of this assumption, the empirical application uses cosponsorship networks from the previous legislative session to explain votes in the current session. The second case corresponds to a scenario where the presence of a cosponsorship link in a congressional session is assumed to be determined solely by observable attributes of the legislators; the existence of the link in a previous congressional session has no impact on whether or not the link is formed in this session. While this assumption initially appears quite restrictive, it represents an important initial step in estimating peer effects in dynamic networks that vary over time. In addition, this assumption is plausible in the context of Congress. Legislators are much more likely form links strategically (i.e. based on observables) than is typical in a friendship network; furthermore, elections regularly change the composition of Congress and make it more plausible that the network resets periodically. The plausibility of these assumptions is a key focus of the empirical application and will be discussed in more detail in Section 1.6.

I begin by exploring a situation where the network does not change and multiple outcomes are observed for each individual in Section 1.2. I generalize the strict exogeneity assumption to allow for individual level unobservables, and I show that with fixed effects estimation identification of peer effects is restored. In Section 1.3 I generalize the framework to examine a single outcome per network from multiple networks with overlapping, but not identical, membership. Estimation in this setting requires assumptions about the network formation process, and I discuss situations where these assumptions are appropriate. I outline a semi-parametric instrumental variables estimation strategy that can estimate these peer effects in Section 1.4. In Section 1.5 I compare the performance of the panel techniques and the semi-parametric method on simulated networks and outcomes. Section 1.6 contains the results of an empirical application to congressional voting patterns. Section 1.7 concludes.

1.2 Peer Effects in a Single Network

Consider the linear-in-means model

$$y_i = \beta_0 + \beta_{\bar{y}}\bar{y}_{(i)} + \beta_x x_i + \beta_{\bar{x}}\bar{x}_{(i)} + \varepsilon_i \tag{1.1}$$

where an individual's outcome is a function of the mean outcome of their peer group ($\bar{y}_{(i)}$), individual level covariates (x), and the mean of peer group covariates ($\bar{x}_{(i)}$).[3] Manski (1993) discusses the estimation of these two kinds of peer effects. The first, an endogenous peer effect ($\beta_{\bar{y}}$), is the effect of the average outcome of an individual's peer group on that individual's outcome. The second, an exogenous peer effect ($\beta_{\bar{x}}$), is the effect of the average of a given attribute of the peer group on an individual's outcome. In addition, Manski draws a distinction between these peer effects and contextual effects, which emerge because individuals in a social group share an environment. As an example to illustrate these different types of effects, consider a situation in which we are trying to estimate peer effects on a student's GPA with family income as the covariate. In this case, the endogenous peer effect is the effect of the student's friends' average GPA on the student's GPA, the exogenous peer effect is the effect of the student's friends' average family income on the student's GPA, and contextual factors are unobservable variables that affect all individuals in the peer group equally, such as teacher quality or school spending.[4]

In Manski's framework, endogenous and exogenous peer effects are unable to be separately identified. As postulated by Moffitt et al. (2001), identification of peer effects fails for three reasons: *simultaneity*, because the outcome enters on both sides of the

---

[3]I follow the notational/expositional convention of Bramoullé et al. (2009) and other literature in this area that assumes only a single covariate x; however, all that follows is general to multiple covariates. The notation I use most closely resembles Goldsmith-Pinkham and Imbens (2013).

[4]Note that this scenario is examined in Goldsmith-Pinkham and Imbens (2013).

regression model, *correlated unobservables*, because all members of a peer group share an environment and an unobservable shock to the environment could be confounded with a peer effect, and *endogenous membership*, which makes it difficult to distinguish between peers being similar to one another because they influence one another and these peers selecting to be friends because they already were similar (this concept is also known as homophily). Bramoullé et al. (2009) proposes a peer effect model that relies on a generic social network structure to identify peer effects and resolves these identification problems, most notably they show that the reflection problem (simultaneity) is resolved because each individual's peer group is different. However, identification requires the assumption that any unobservable attributes that affect the outcome do not affect the network formation process; if this assumption is violated peer effects are not identified and estimation of model parameters is inconsistent. I introduce the idea of social networks as peer groups below and show that a generalization of the peer effects model can account for individual level unobservables, which restores identification and consistent estimation.

1.2.1 Social Networks as Peer Groups

Peer effects are traditionally estimated when individuals are partitioned into distinct groups such as classrooms or offices. Social networks allow for estimation of peer effects based on peer groups that are different for each individual. Consider a generic

social network described by an adjacency matrix A, where A_{ij} represents the relationship between individuals i and j. It is useful to have the matrix in row-normalized[5] form G, where each of the links between individuals are weighted by the total number of links $(\sum_{j=1}^{n} A_{ij} = M_i)$ that an individual has, i.e.

$$G_{ij} = \frac{A_{ij}}{\sum_{j=1}^{n} A_{ij}} = \frac{A_{ij}}{M_i}. \tag{1.2}$$

Adjacency matrices are quite flexible, allowing for directed relationships, where an individual influences another without being influenced back $(A_{ij} \neq A_{ji})$, and weighted relationships, where some links are allowed to be more important. Because of this flexibility, a social network nests previous peer effects models, including the formulations of Manski (1993), Moffitt et al. (2001), and Lee (2007).[6] In all of these cases the adjacency matrix consists of block diagonal peer groups where all individuals in a given group are weighted equally.

Returning to the linear in means model in Equation 1.1, it is possible to express the group means as function of the network. The endogenous peer effect is a weighted sum

---

[5] A row-normalized matrix has the advantage of scaling elements of the adjacency matrix so that the strength of relationships can be compared across individuals; for example in the application this takes into account the frequency with which legislators cosponsor bills. Additionally, Bramoullé et al. (2009) and Lee (2007) note that a row-normalized matrix is more likely to meet the assumptions required for identification of the peer effects. Note that calculating the row-normalized form assumes that there are no isolated nodes (i.e. nodes that have no links) or that isolated nodes are removed from the network.

[6] Bramoullé et al. (2009) demonstrate that these formulations share their original identification properties.

of of an individual's peers as given by the adjacency matrix:

$$\bar{y}_{(i)} = \frac{1}{M_i} \sum_{j=1}^{n} A_{ij} y_j = \sum_{j=1}^{n} G_{ij} y_j \tag{1.3}$$

where n is the number of individuals in the network and M_i is the number of friends

individual i has (the number of nonzero entries on row i of A). The mean for the

exogenous variable $\bar{x}_{(i)}$ is defined similarly. By convention the diagonal of the adjacency

matrix is set to 0, ruling out the possibility of self-links, which results in i's own outcome

being omitted from the group mean calculation. In this sense the model resembles

the formulation of Moffitt et al. (2001), which calculates group means omitting the

individual in question. Additionally, it is useful to write the model in matrix form,

$$\mathbf{y} = \beta_0 \mathbf{i} + \beta_{\bar{y}} G \mathbf{y} + \beta_x \mathbf{x} + \beta_{\bar{x}} G \mathbf{x} + \varepsilon \tag{1.4}$$

where $\mathbf{i}$ represents a vector of ones of length n. Note that as a direct result of matrix

multiplication and row-normalization, $G\mathbf{y}$ results in a vector of length n where each

entry is $\bar{y}_{(i)}$.

Bramoullé et al. (2009) establish conditions for separate identification of the exoge-

nous and endogenous peer effect under the assumption of strict exogeneity $E[\varepsilon|\mathbf{x}, G] = 0$.

In addition to the exogeneity assumption, identification requires that the matrices I_n, G,

and G^2 are linearly independent.[7] Notably, the strict exogeneity assumption is violated

---

[7]Intuition for the linear independence condition will be discussed below in the generalized model. Identification of the reduced form also requires the condition that $\beta_x \beta_{\bar{y}} + \beta_{\bar{x}} \neq 0$; however, a violation of this condition is so unlikely to occur that it will be ignored through the rest of the paper.

if there are unobserved factors that are correlated with both the outcome of the model

and the network formation process, i.e. the assumption requires that the only factors that

impact the network formation process are those that are also included in the regression

model. This assumption is unlikely to hold in many settings, including the application in

Bramoullé et al. (2009) who attempt to find peer effects in the number of extracurricular

activities in which students participate. If a student has a fixed unobservable attribute,

such as extroversion, that affects both the number of friends they have (or who they make

friends with) and the number of extracurricular activities they pursue, strict exogeneity

does not hold. Generalizing this assumption to correct for these unobservable attributes

is the focus of the remainder of the section.

It is important to note that this framework does not require that the social groups

be imposed exogenously.[8] The network can be endogenously determined, in the sense

that peer groups can be the result of individual choices, but these choices must be

uncorrelated with the individual error term ε_i. Otherwise, we risk confounding an

individual peer effect with effects that actually result from sorting on unobservables.

Stated another way, the data generating process for the outcome is assumed to be

independent from the process that formed the network. This impacts the subsequent

results in two distinct ways. I initially begin by assuming that the network is fixed when

---

[8]While these assumptions allow estimation of the model with endogenous social groups, imposing exogenous groups often means the assumptions required for identification are met ex ante. Thus, the common practice in economics of relying on exogenous variation in group structure to identify peer effects is nested as a special case, whereas the work here emphasizes applications to endogenously formed networks and discusses issues that arise when examining these networks.

the outcomes are generated. Subsequently, I allow networks that vary across outcomes

and I discuss the assumptions about the network generation process that are required

for this to be valid. Mathematically the two assumptions are both formalized by taking

expected values conditional on the specific realization of the network generated by some

network generation process. In addition, I proceed assuming that the observed network

perfectly represents the underlying social structure. Sampled networks or networks that

are measured with error can have significant and somewhat unpredictable effects in peer

effects estimation (see Chandrasekhar and Lewis, 2012).

1.2.2 Fixed Network with Multiple Outcomes

Consider a variation of the peer effects model where the outcome *y* and the exogenous

variable *x* are observed *B* times for each individual and the network does not change.[9]

The regression model then becomes

$$y_{ib} = \beta_0 + \beta_{\bar{y}} \bar{y}_{(i)b} + \beta_x x_{ib} + \beta_{\bar{x}} \bar{x}_{(i)b} + \varepsilon_{ib} \tag{1.5}$$

where *b* indexes the observed outcomes and exogenous variables. Now that the model

includes multiple observations for each individual I make the following assumption:

Assumption 1 (Individual Heterogeneity)**.** *The error term takes the form* $\varepsilon_{ib} = \gamma_i + \nu_{ib}$

---

[9]This situation is not as uncommon as it may seem and is discussed in Section 1.7. For now, suffice it
to observe that this corresponds to an experiment where people are divided (or select) into groups and
multiple outcomes are observed, possibly over time. As formulated this model also has many similarities
to models derived in geography and urban economics. In fact, the estimation strategy proposed by
Bramoullé et al. (2009) is built upon techniques from spatial econometrics.

Assumption 1 decomposes the error term into a fixed attribute γ_i that does not vary across outcomes and an idiosyncratic term v_{ib}. I proceed with the conventional assumption that γ_i and v_{ib} are uncorrelated which is necessary for identification. Explicitly including the individual fixed effect γ_i in the regression model in Equation 1.5 yields:

$$y_{ib} = \beta_0 + \beta_{\bar{y}}\bar{y}_{(i)b} + \beta_x x_{ib} + \beta_{\bar{x}}\bar{x}_{(i)b} + \gamma_i + v_{ib} \tag{1.6}$$

and now it is straightforward to take the conventional within transformation[10] to account for the individual fixed effect

$$\left(y_{ib} - \frac{\sum_{b=1}^{B} y_{ib}}{B}\right) = \beta_{\bar{y}}\left(\bar{y}_{(i)b} - \frac{\sum_{b=1}^{B}\bar{y}_{(i)b}}{B}\right) + \beta_x\left(x_{ib} - \frac{\sum_{b=1}^{B} x_{ib}}{B}\right)$$
$$+ \beta_{\bar{x}}\left(\bar{x}_{(i)b} - \frac{\sum_{b=1}^{B}\bar{x}_{(i)b}}{B}\right) + \left(v_{ib}\frac{\sum_{b=1}^{B} v_{ib}}{B}\right)$$

where each term in the regression is de-meaned by its average over all observed outcomes. This model is written more concisely as

$$y_{ib}^* = \beta_{\bar{y}}\bar{y}_{(i)b}^* + \beta_x x_{ib}^* + \beta_{\bar{x}}\bar{x}_{(i)b}^* + v_{ib}^* \tag{1.7}$$

where the stars denote that the variable has undergone a within transformation. As before, this model can also be expressed in matrix form

$$\mathbf{y}_b^* = \beta_{\bar{y}}G\mathbf{y}_b^* + \beta_x\mathbf{x}_b^* + \beta_{\bar{x}}^*G\mathbf{x}_b^* + \mathbf{v}_b^* \tag{1.8}$$

---

[10]As is frequently the case in these situations, this process can also be solved through explicit estimation of fixed/random effects parameters. I focus on the within transformation primarily to limit the number of parameters that must be estimated as the Bayesian estimation strategies used in this paper can be computationally intensive; however, in many settings (especially short panels) it may be desirable to specify the model differently. See also Lee and Yu (2010) and Parent and LeSage (2012) for discussions of fixed and random effects in spatial models that address similar issues.

under the assumption that the network is fixed for all outcomes. The model can also be

expressed over all outcomes as

$$\tilde{\mathbf{y}}^* = \beta_{\tilde{y}}\mathbf{G}\tilde{\mathbf{y}}^* + \beta_x\tilde{\mathbf{x}}^* + \beta_{\tilde{x}}\mathbf{G}\tilde{\mathbf{x}}^* + \tilde{\nu} \tag{1.9}$$

where $\tilde{\mathbf{x}}^*, \tilde{\mathbf{y}}^*$ and $\tilde{\nu}^*$ are vectors with dimension $(nB \times 1)$ and the matrix $\mathbf{G} = G \otimes I_b$.[11]

Once the model has been transformed in this fashion, identifying peer effects requires

the following exogeneity assumption:

Assumption 2 (Strict Exogeneity with Individual Heterogeneity)**.** $E[\nu_b^*|\gamma, \mathbf{x}_b^*, G] = 0$

where γ is a vector of individual level fixed effects. Generalizing the exogeneity assump-

tion in this fashion makes it much more plausible that the network formation process

meets the conditional exogeneity assumption; fixed individual unobservable attributes

are now assumed to affect the formation of network links in addition to the outcome.

Note that this assumption implies that the error term is uncorrelated across outcomes,

ruling out any learning behavior or time-series attributes of the outcomes.[12]

Once the model is transformed to account for individual heterogeneity peer effects

are identified; formally stating the result:

---

[11]$\mathbf{G}$ is a $nB \times nB$ block diagonal matrix with G repeated along the diagonal B times.

[12]Parent and LeSage (2012) provides a model that more directly addresses temporal dependence and random effects in spatial econometric models that provide insight into how a peer effects model might account for these effects if necessary.

Proposition 1.1. *Under Assumption 1 and Assumption 2, peer effects are identified in the multiple outcomes peer effects model in Equation 1.9 if the matrices I_{nB}, $\mathbf{G}$, and $\mathbf{G}^2$ are linearly independent.*

Proof. The transformed model, combined with the new strict exogeneity assumption and the linear independence restriction, meets the conditions for identification established by Bramoullé et al. (2009), proposition (1). □

Note that the linear independence restriction is required to find an identifying instrument to estimate the model. If this condition is not met, there are no valid instruments and the parameters cannot be identified. The intuition behind this condition, along with details of the estimation procedure, is explained in Section 1.4.

While the within transformation accounts for individual heterogeneity, it is also possible to account for outcome level unobservables that potentially confound estimation (analogous to time-period fixed effects). This loosens the restriction that the outcome not exhibit time-series behavior by accounting for unobservable outcome level attributes that affect all individuals in the same way. Accounting for outcome level unobservables requires the following modification of Assumption 1,

$$\varepsilon_{ib} = \gamma_i + \delta_b + v_{ib}$$

where the error term is decomposed into an individual fixed effect γ_i an outcome level effect δ_b and an idiosyncratic error term v_{ib}. In addition, the exogeneity assumption is

now conditioned on the outcome level fixed effect δ

$$E[v^*|\gamma, \delta, \mathbf{x}^*, \mathbf{G}] = 0.$$

Similar to the above result, peer effects are identified in this model because it fits the assumptions outlined in Bramoullé et al. (2009), proposition (4).

In this section I have shown that peer effects are identified even in the presence of individual heterogeneity. In contrast to a single network, where identification fails unless all individuals have an error term with zero-conditional mean, when the network is held constant and multiple outcomes are observed, fixed effects estimation can restore identification by accounting for time invariant individual unobservable characteristics. I now turn to a scenario where networks are allowed to change over time.

1.3   Peer Effects Across Multiple Networks

An alternate way to account for individual heterogeneity is to examine a framework with one outcome in each of multiple networks rather than multiple outcomes from a single network, adding the requirement that individuals belong to more than one network. This contrasts with Bramoullé et al. (2009) who generalize their model across multiple networks but proceed with the assumption that these networks are part of a single cross-section and are completely independent from one another. This section shows how peer effects can still be identified as long as the network formation process meets certain

assumptions, namely link formation must be random conditional on observable attributes

of the individuals and the only correlations between separate networks are due to these

attributes.

1.3.1 Bramoullé Multiple Network Model

Consider the following generalized peer effects model where data is gathered from

multiple networks, indexed by ℓ, and peer groups are potentially endogenous as before:

$$y_{i\ell} = \beta_{0\ell} + \beta_{\bar{y}}\bar{y}_{(i)\ell} + \beta_x x_{i\ell} + \beta_{\bar{x}}\bar{x}_{(i)\ell} + \varepsilon_{i\ell} \tag{1.10}$$

where i indexes the observations in the data and y_i is observed only once in a network.

The means of the peer groups are defined as before for each individual network so that

$$\bar{y}_{(i)\ell} = \frac{1}{M_i}\sum_{j=1}^{n_\ell} A_{ij\ell}y_{i\ell} = \sum_{j=1}^{n_\ell} G_{ij\ell}y_{i\ell} \tag{1.11}$$

where n_ℓ is the number of individuals in network ℓ. The peer effects model can be

transformed as before under the assumption that the error term can be decomposed as

Assumption 3 (Individual Heterogeneity in Multiple Networks). $\varepsilon_{i\ell} = \gamma_i + \nu_{i\ell}$

and written in matrix form for a given network G_ℓ

$$\mathbf{y}_\ell^* = \beta_{\bar{y}}G_\ell\mathbf{y}_\ell^* + \beta_x\mathbf{x}_\ell^* + \beta_{\bar{x}}G_\ell\mathbf{x}_\ell^* + \gamma + \nu_\ell. \tag{1.12}$$

As in Section 1.2, the error term ν_ℓ is assumed to be independent of the individual

level attribute γ, the exogenous variable $\mathbf{x}_\ell$, and the network G_ℓ,

Assumption 4 (Strict Exogeneity in Multiple Networks). $E[\nu_\ell | \gamma, \mathbf{x}_\ell, G_\ell] = 0$.

This strict exogeneity condition implies that the error term is uncorrelated with the process that generated the network *G for each network*; however, this does not imply that the network is exogenously determined. The network formation process can be a function of individual level observables as long as they are accounted for in the model.

In order to estimate peer effects across multiple networks it is necessary to assume that each network is independent from one another, conditional on exogenous variables and time invariant unobservable attributes. This allows the formation of $\mathbf{G}_\ell$, a block diagonal matrix where the diagonal consists of matrices G_ℓ.[13] If this assumption is satisfied, along with Assumption 3 and Assumption 4, peer effects are identified as before:

Proposition 1.2. *Under Assumption 3 and Assumption 4, peer effects are identified in the multiple networks peer effects model in Equation 1.12 if the matrices I_N, $\mathbf{G}_\ell$, and $\mathbf{G}_\ell^2$ are linearly independent where I_N is an identity matrix of dimension $N = \sum_{\ell=1}^{L} n_\ell$.*

Proof. Similar to the previous results, the transformed model, the new strict exogeneity assumption, and the linear independence restriction meet the conditions for identification established by Bramoullé et al. (2009), proposition (1). □

---

[13]If this assumption is violated the off-diagonal elements will not be zero. For example, the lagged value of G_{ij} will have some effect on G_{ij}.

In addition, note that if peer effects are identified in each individual network in the joint model, peer effects will be identified across all networks; however, the converse is not true. Peer effects are not necessarily identified in an individual network even if they are identified in the model with all networks. These results follow directly from the requirements for identification in peer effects and are shown in Section A.1.

1.3.2 Accounting for Contextual Effects

In contrast to the previous section, where the network is held fixed, analyzing multiple networks requires accounting for contextual effects, environmental factors that affect all individuals in the network. These are analagous to the outcome fixed-effects in the previous section, but in this case represent a specific challenge to identification of peer effects. In the prior model with a fixed network, there is arguably a single "context" for all of the outcomes; whereas, multiple networks each potentially have their own contextual factors that must be accounted for in the estimation procedure.

Contextual effects are accounted for using fixed effects which in this case has a specific matrix format. Since all individuals share the same contextual factors, we can subtract the average of an individual's friends' variables from individual i's variables. Note that the average outcome of an individual's friends is given by $G_\ell \mathbf{y}_\ell$ and subtracting this quantity from the outcome yields $\mathbf{y}_i - G_\ell \mathbf{y}_\ell = (I - G_\ell)\mathbf{y}_\ell$. In matrix form, the model

in Equation 1.12 becomes

$$(I_{n_\ell} - G_\ell)\mathbf{y}_\ell = \beta_{\bar{y}}(I_{n_\ell} - G_\ell)G_\ell\mathbf{y}_\ell + \beta_x(I_{n_\ell} - G_\ell)\mathbf{x}_\ell + \beta_{\bar{x}}(I_{n_\ell} - G_\ell)G_\ell\mathbf{x}_\ell + (I_{n_\ell} - G_\ell)\varepsilon_\ell$$

(1.13)

where I_{n_ℓ} is an $n_\ell \times n_\ell$ identity matrix. Bramoullé et al. (2009) refer to this as the "within local" transformation and show that peer effects are identified in this model if the matrices, I_{n_ℓ}, G_ℓ, G_ℓ^2, and G_ℓ^3 are linearly independent, with the additional term required because the entire model has been multiplied by the matrix G_ℓ. While it is true that the within-local transformation will account for the contextual effect, it differs from the conventional fixed effects transformation which subtracts the average of all individuals (as opposed to just friends) from an individual's outcome. Define the matrix H as $H_\ell = \frac{1}{n_\ell}\mathbf{ii}'$, then the "within-global" transformation is given by

$$(I_{n_\ell} - H_\ell)\mathbf{y}_\ell = \beta_{\bar{y}}(I_{n_\ell} - H_\ell)G_\ell\mathbf{y}_\ell + \beta_x(I_{n_\ell} - H_\ell)\mathbf{x}_\ell + \beta_{\bar{x}}(I_{n_\ell} - H_\ell)G_\ell\mathbf{x}_\ell + (I_{n_\ell} - H_\ell)\varepsilon_\ell$$

(1.14)

In this case, peer effects can be identified if the $rank(I_{n_\ell} - G_\ell) < n - 1$ even if the matrices above are linearly dependent.[14] For this reason Bramoullé et al. (2009) state that this model has the least restrictive identification assumptions, and I implement a global within transformation when necessary in estimating peer effects.

---

[14]Note that if peer effects are identified with the within-local transformation they will necessarily be identified in a within-global transformation. This new identification condition outlines scenarios where the within-global transformation provides identification even when the within-local does not and is especially salient for directed networks. For example, Bramoullé et al. (2009) show a case where the matrices G and G^3 are linearly dependent, but because the network is directed the rank condition still holds.

1.3.3 Multiple Observations across Multiple Networks

It is possible to combine the above results with results from Section 1.2 to examine

multiple outcomes in multiple networks. For each observed network ℓ with multiple

outcomes B_ℓ the generalized version of Equation 1.9 is

$$\tilde{\mathbf{y}}_\ell^* = \beta_{\bar{y}} \tilde{G} \tilde{\mathbf{y}}_\ell^* + \beta_x \tilde{\mathbf{x}}_\ell^* + \beta_{\bar{x}} \tilde{G} \tilde{\mathbf{x}}_\ell^* + \tilde{\nu}_\ell \tag{1.15}$$

By the results in Section 1.2 and Section 1.3, as long as peer effects are identified in each

network, they will be identified in the whole model (see Section A.1). This model can

account for individual level unobservables, outcome level unobservables, and network

level unobservables through the various transformations discussed up to this point.

While I have shown that peer effects can be identified in multiple networks over

the same set of individuals, it is important to note that this setting differs from that a

single social network that is allowed to evolve overtime. The techniques derived in this

section correspond to a scenario where each new network is independently generated

(or can be assumed to be independent). Examining a single network that evolves over

time will require a different methods because in this setting links are dependent; a link

that appears in one network is highly likely to appear in subsequent networks.[15] It may

---

[15] An additional complication that would need to be addressed in a dynamic setting is that it is highly likely that subsequent links may be formed based on prior outcomes. This issue is discussed in Goldsmith-Pinkham and Imbens (2013) who examine two networks sampled at different times from the same high school.

be possible to generalize these peer effects models to allow for some form of network serial correlation; however, this task is left for future research.

1.4 Estimation Strategy

Because the exogenous peer effect is an endogenous right hand side variable, the model must be estimated using instrumental variables. Consider as an example the model in Equation 1.12 where there are multiple networks with a single observation for each network. The first stage regression of the instrumental variables model is used to estimate the endogenous peer effect $\mathbf{G}_\ell \mathbf{y}_\ell^*$ and is given by

$$\mathbf{G}_\ell \mathbf{y}_\ell^* = \alpha_x \mathbf{x}_\ell^* + \alpha_{\bar{x}} G_\ell^* \mathbf{x}_\ell^* + \alpha_{IV} (G_\ell^*)^2 \mathbf{x}_\ell^* + \eta_\ell^* \tag{1.16}$$

where the instrument is $(G_\ell)^2 \mathbf{x}_\ell^*$. The square of an network adjacency matrix identifies the friends of an individuals friends. Building on previous literature in spatial econometrics, Lee (2007) suggests that higher order polynomials of the network multiplied by exogenous covariates (i.e. $G^2 x$, $G^3 x$ etc.) are valid instruments. Because each individual's outcome is a function of their own covariates and their friends' covariates, the endogenous peer effect (the effect of peer group outcomes, $\mathbf{G}_\ell \mathbf{y}_\ell^*$) is a function of the covariates of the friends of an individual's friends, $(G_\ell)^2 \mathbf{x}_\ell^*$.

In addition, the instrument is uncorrelated with the error term. Because of the strict exogeneity assumption in Assumption 4, the instrument is a function of the observed

network and the observed exogenous variable $\mathbf{x}^*$ and is therefore uncorrelated with

the error term.[16] Note that the model specification assumption has a crucial role here,

as it implies that the only variables that systematically affect an individual's outcome

are those explicitly included in the model. In other words, it is assumed that there

are no direct effects of friends of friends, rather the only way that friends of friends

can influence the outcome is through the mutual friend.[17] Because the instrument is

uncorrelated with the error term and correlated with the endogenous variable it is a valid

instrument.

I implement the instrumental variables model outlined above in two distinct and novel

ways. The first is to use the within transformations from Section 1.2 and Section 1.3

to account for individual heterogeneity and then estimate the instrumental variables

model. I do this using a Markov-Chain Monte Carlo instrumental variables estimation

procedure. The second is to estimate the model using a semi-parametric technique.

This method, proposed by Conley et al. (2008), relies on a non-parametric technique to

group observations that are likely to share a distribution for their error term. Once the

grouping is known, the parametric instrumental variables estimation proceeds using this

knowledge. Each group of observations has its own mean and variance parameters, which

means that individual heterogeneity is flexibly accounted for in the model, potentially

---

[16]By similar logic, higher order powers are also valid instruments. This could allow modeling a specific effect for friends of friends.

[17]Bramoullé et al. (2009) refer to these kinds of relationships as intransitive triads, and examine through simulation the identification properties of the model as the number of these triads change.

even without directly accounting for an individual fixed effect. Both methods are described in more detail below.

1.4.1 Instrumental Variables and Gibbs Sampling

Prior research in estimating these instrumental variables models suggests using a Maximum Likelihood approach, with Lee (2007) observing that even though a conventional two-stage least squares estimate is consistent, in this setting its rate of convergence is far slower than would otherwise be expected.[18] Unfortunately, the properties of these maximum likelihood estimators are still not very well understood. Specifically, while Lee (2007) shows that these estimators are approximately quadratic at the optimum, no large-sample results have been established for statistical inference using parameters and standard errors derived from the likelihood maximization. In addition, the likelihood functions in this case can be very difficult to optimize and it is unclear whether they always converge to a global maximum. Goldsmith-Pinkham and Imbens (2013) acknowledge the limitations of maximum-likelihood estimation and introduce a Bayesian estimation strategy to estimate the reduced form of the linear-in-means model; the method performs well, but is only implemented with a single outcome.

I pursue a Bayesian instrumental variables estimation strategy for two reasons. First, Bayesian methods have well-established optimality results and avoid the pit-

---

[18]See also Lee et al. (2010). While not an application of social networks, Lin (2010) applies the Lee (2007) model to identify peer effects in schools.

falls in maximum likelihood estimation outlined above. Second, Bayesian methods result in estimating posterior distributions for parameters that allow clear estimation of confidence intervals and interpretation of parameters. Finally, the semi-parametric estimation strategy I propose to resolve individual heterogeneity is a Bayesian model where maximum-likelihood estimation is unfeasible. Estimating the panel approaches to the peer effects models as a Bayesian instrumental variables allows more direct comparison of the methods. Of course, Bayesian methods are not without limitations. Specifically, estimation can be computationally intensive and requires assumptions of prior distributions, which some researchers are reluctant to do.

Estimating the models in this paper via maximum likelihood or via Bayesian Markov-Chain Monte Carlo methods requires a normality assumption.[19] In this case, the error terms for the first and second stage are assumed to be jointly normal with mean zero and variance Σ

Assumption 5 (Normality of the Error Terms)**.**

$$v, \eta \,|\, x^*, G \sim N(0, \Sigma)$$

Note that this assumption is not necessary for identification, which will hold as long the exogeneity assumption is satisfied. This allows modifications of this model that vary

---

[19]It is worth noting that the Bernstein - Von Mises theorem implies that for a sufficiently large sample the maximum likelihood and Bayesian estimates are equivalent. Consequently, if one is more comfortable with a MLE approach, this intuition should suffice to interpret the estimates even if one is unfamiliar with Bayesian techniques.

this distribution assumption while maintaining identification. For example, this model can be used to estimate binary outcomes via a probit or logit with minimal changes to the assumptions used in the Bayesian estimation model.

I fit the instrumental variables model via Gibbs sampling, a common Bayesian Markov-Chain Monte Carlo (MCMC) technique. MCMC uses an iterative process to estimate posterior distributions for the model parameters. Implementing this method requires the assumption of prior distributions; in this case, the model parameters are assumed to be normally distributed normally distributed with mean zero and a sufficiently large variance to be considered a diffuse prior. The error term Σ is assumed to have an Inverse-Wishart distribution. Details of these assumed distributions are specified for the models as implemented and explained in greater depth in the online appendix.

1.4.2 Semi-Parametric Estimation

There are many scenarios where using a within transformation to account for individual heterogeneity is unfeasible. For example, panel data may not be available or there may be concerns about time-varying individual heterogeneity. To address these concerns, I suggest using a semi-parametric instrumental variables estimator suggested by Conley et al. (2008). Estimating the model parameters is done using a very similar method to the Gibbs sampling, with the exception that each observation is assumed to have its own mean and variance so that

Assumption 6 (Normality of the Error Terms: Dirichlet Process)**.**

$$v_i, \eta_i | x^*, G \sim N(\mu_i, \Sigma_i)$$

To avoid estimating more parameters than there are observations, there must be some limitations placed on the number of these parameters. In this case the distribution parameters μ_i and Σ_i are assumed to depend on the group assignment J of the observation, where each observation is assigned to a specific group. The group assignments are determined endogenously via a Dirichlet Process (DP)[20], $J \sim DP(\lambda, J_O)$ where λ and J_0 are the hyperparameters for the Dirichlet Process.[21]

The Dirichlet Process prior results in a mixture of normals, where the error term is normally distributed but the exact distribution is determined by what group the observation belongs to. This is done non-parametrically, through the Dirichlet Process, while the models are traditional linear instrumental variables; hence, this technique is semi-parametric. Note that it includes an individual mean and variance so that the estimation process can directly account for individual level unobservables, no matter their source. This semi-parametric technique is quite powerful, in that it requires no ex ante assumptions about the nature of the heterogeneity beyond assuming that each group shares a distribution family (in this case normal distribution). The process of choosing the most likely grouping of observations naturally determines what attributes

---

[20]Also known as a Chinese Restaurant Process, this is a statistical model of group formation that has several useful properties for Bayesian estimation.

[21]The prior distribution assumptions for these parameters are somewhat involved and are covered in great detail in Conley et al. (2008); I implement this model exactly as outlined in their paper.

of the observations most differentiate them one from another. For example, there may be a situation where we are concerned both about individual and time period fixed effects; however, there is insufficient data to account for both and still have enough degrees of freedom to conduct inference. A Dirichlet Process can group observations so that the factors that most differentiate observations from one another are accounted for, while those that are not relevant can be ignored.[22]

In addition to allowing for individual level means, Conley et al. (2008) show that this semi-parametric model has several useful features. Because it assumes each observation has its own variance, it explicitly accounts for heteroskedasticity. Partly due to this, simulations reveal it to be at least as efficient as many other instrumental variables estimation techniques, including both Frequentist and Bayesian methods. Additionally, a Dirichlet Process has natural conjugate priors so that estimation can proceed efficiently via Gibbs sampling. This is especially true when the error is non-normal. Lastly, the Dirichlet Process technique is demonstrated to behave well even in the case of weak instruments.

The Dirichlet process is not without limitations. The algorithm finds groups based on the observed data and these endogenously discovered groups may or may not reflect the attributes that a researcher considers important. In a related concern, the algorithm may not be able to group observations at a fine enough level to completely account for all

---

[22]For example, the Dirichlet Process might find that rather than an individual level mean it is sufficient to account for a few group means and that all time fixed effects are unnecessary.

the individual heterogeneity when observations are substantially different. In addition, the semi-parametric model can be computationally intensive and may not behave well in large datasets. There are two reasons for this. The first is that as more observations are added to the model, sorting into groups becomes increasingly difficult. Additionally, each group of observations requires its own draw of mean and variance parameters, so as the number of groups increases additional computations are required within a given iteration of the algorithm. The combination of these two problems means that the semi-parametric method does not scale well.

Despite these limitations, the Dirichlet process semi-parametric estimation strategy can still be a powerful tool to estimate these models due to its flexibility. Consequently, I suggest that fixed effects transformation and semi-parametric estimation are complementary techniques, the choice of which model to estimate is likely case specific. In most cases, estimating both as a robustness check seems prudent. In addition, there is no reason why the methods cannot be used in tandem when the data is available. The fixed effects model can directly account for individual fixed effects while the semi-parametric estimation strategy can flexibly account for any remaining heterogeneity in the observations.

1.5 Simulations

To further explore the attributes of these estimation procedures I estimate them on a variety of simulated networks and data to observe how they behave. The generic process consists of generating a random network, simulating data generated from that network, and then fitting a model using a variety of the estimation techniques suggested above. Unless otherwise stated, outcomes are simulated 1000 times and the outcome of interest is the endogenous peer effect. While there are many potential ways to simulate network data, I focus on two network simulation models.

The first, Erdös-Rényi random graphs, is the most common and likely simplest method for simulating a network. In an Erdös-Rényi random graph a link is formed between two individuals as a draw from a Bernoulli distribution with a given link formation probability. This means that the network can be generated very simply; however, the resulting networks differ somewhat from social networks that are observed in the real world.[23] These networks are commonly used in the literature to benchmark the performance of network techniques, and I use this method to ensure that the techniques suggested in this paper perform comparably to the original method.

---

[23]Specifically, Erdös-Rényi random graphs lack the clustering that is commonly observed in real world networks and the degree distribution (the distribution of the number of links each individual has) is uniform, rather than the skewed "power law" distributions observed in many networks.

Table 1.1: Random Network Simulations, Erdös-Rényi

	Coverage	SD	Converge
IV	0.73	0.17	0.79
IV - Individual FE	0.98	0.02	0.96
IV - Ind/Net FE	0.99	0.03	0.90
Semiparametric	0.96	0.04	0.82
Semiparametric - Individual FE	0.99	0.01	0.92
Semiparametric - Ind/Net FE	1.00	0.03	0.93

The second random network generation technique, known as Exponential Random Graph Models (ERGMs), are able to generate networks that are much more realistic.[24] While the main focus of ERGMs is to statistically evaluate the likelihood of observing a given network, in doing so they establish methods for simulating networks that have certain attributes. Specifically, ERGMs can generate networks that exhibit homophily (selection on observables) and allow each individual to have an arbitrary degree distribution (a different number of expected links).

In Table 1.1 I show the simulated results from the scenario outlined in Section 1.2 where the network is held fixed and multiple outcomes are observed for each network. In this case the networks are Erdös-Rényi random graphs. The top panel contains results for the Gibbs sampling, where each row varies the kind of fixed effects that are accounted for in each model, and the bottom panel contains results from the Dirichlet Process

---

[24]There is a large and growing literature examining the properties of ERGMs. Most of these focus on either fitting ERGMs to observed networks or establishing properties of the estimation procedure. Chandrasekhar and Jackson (n.d.) provide a high level introduction into these models and some of their limitations.

Table 1.2: Random Network Simulations, ERGM with Homophily

	Coverage	SD	Converge
IV	0.63	0.07	0.75
IV - Individual FE	0.98	0.01	0.96
IV - Ind/Net FE	0.99	0.02	0.91
Semiparametric	1.00	0.03	0.81
Semiparametric - Individual FE	0.99	0.01	0.95
Semiparametric - Ind/Net FE	0.99	0.02	0.96

semi-parametric estimation procedure. Each group of columns specifies a different kind of data generating process, varying the kinds of individual heterogeneity in the model. The estimation procedures are evaluated on whether or not the final 95% highest probability density (HPD) interval covers the true parameter (Coverage), the median standard deviation for the estimated parameter (Precision), and whether or not the Bayesian estimation procedure converged (Converge).[25] The results of the table indicate that both techniques are capable of estimating the correct parameters, with the Gibbs sampling requiring the fixed effects transformations to accurately estimate the model. As expected, the semi-parametric estimation strategy appears to function regardless of whether or not the data has been transformed, although it is more likely to converge in the transformed model.[26]

---

[25]Convergence is tested using a Geweke Test with 95% confidence. While it is expected that most of these algorithms will have converged in distribution, due to the nature of the statistical test it is possible to reject convergence simply due to noise in the estimated output. In other words, I expect the Convergence diagnostic to indicate convergence roughly 95% of the time.

[26]This problem is due to the untransformed model having many additional groups to find in the Dirichlet Process. Theoretically, this can be remedied in most cases by allowing the algorithm to run for more iterations; however, in these tables I wanted to compare methods holding the number of iterations constant.

Table 1.3: Random Network Simulations, ERGM with Unobserved Homophily

	Coverage	Precision	Converge
IV	0.55	0.13	0.75
IV - Individual FE	0.90	0.01	0.93
IV - Ind/Net FE	0.90	0.03	0.93
Semiparametric	0.90	0.03	0.85
Semiparametric - Individual FE	0.92	0.01	0.95
Semiparametric - Ind/Net FE	0.92	0.02	0.91

I also implement this model when the network is generated by an ERGM that allows for homophily. In this case, I assume that students are more likely to form links when the value of their unobservable individual fixed effect is similar. This corresponds to a case where the network is generated by something unobservable to the modeler, but is accounted for in the estimation procedure by a fixed effect. Table 1.2 shows that the estimation procedures are indeed able to correctly estimate the parameters of the model even when the network is more sophisticated. Table 1.3 shows estimation when the network exhibits sorting on some other attribute that is not part of the data generating process. In this case as well, the parameters of the model are accurately estimated. Taken together, these results suggest that the estimation procedures behave as expected and are robust to several kinds of networks.

Additionally, I am able to fit an ERGM to the political networks in the application to get a better idea of how the estimation procedures will behave in scenarios similar to the empirical application. This is especially important given that congressional

cosponsorship networks have attributes that Bramoullé et al. (2009) indicate will make identifying peer effects more difficult, namely it is densely connected and the diameter is small.[27] Despite these attributes, simulations indicate that peer effects are precisely identified in networks similar to the ones under consideration in the model.

Several conclusions are evident from the simulation results:

- Estimation is successful for a variety of network types

- Identification seems to hold even in networks that exhibit sorting on unobservables

- Semi-parametric estimation can account for individual heterogeneity in a single network

- A hybrid approach yields most precise estimates in many situations

1.6 Application

Congressional voting is an ideal place to look for peer effects. Beginning with Peltzman (1984, 1985) there is a long history in economics of using econometrics to explain voting patterns (see also Levitt, 1996). However, the simultaneity of the voting process has always made it difficult to study the peer effects that emerge in the voting process. I use cosponsorship networks, combined with the results in this paper, to address this question. Bill cosponsorship decisions provide a social network, because a decision to cosponsor provides a clear indication that the cosponsor holds the bill and/or the bill's

---

[27]The diameter is longest direct path between two individuals in a network.

sponsor in high regard. The purpose of this application is to examine the properties of the two estimation procedures outlined above on observed networks and estimate peer effects in congressional networks. Specifically, I wish to see if the relationships indicated by cosponsorship result in peer effects in voting, i.e. do a legislator's friends have a significant influence on their voting decisions?

There has been some research that attempts to estimate peer effects from a legislator's social interactions. These studies typically exploit some sort of exogenous variation in whom legislators may associate with, for example Masket (2008) uses random assignment of seats in the California assembly to show significant peer effects in an individual's voting outcomes based on the voting of those who sit next to them. However, on a national level Rogowski and Sinclair (2012) are unable to identify any peer effects in voting outcomes based on office location, which is partially determined by random assignment.

1.6.1 Cosponsorship Networks

This section introduces the networks and data that will be used as an illustration of the methods described above. Fowler (2006a,b) proposes and explores the idea that bill cosponsorship represents a social network over the set of legislators in the House and Senate. Specifically, Fowler (2006a) suggests that one formulation of the cosponsorship network best describes the underlying social structure of Congress. This network is

directed (links do not have to be mutual); a cosponsor i links to a sponsor j when they choose to cosponsor the bill that j is sponsoring. Additionally, these networks are weighted in a specific fashion. One quirk of the legislative process is that some bills have many cosponsors, occasionally well over half of the membership.[28] It is reasonable to assume in these cases that cosponsorship is a less meaningful connection, both because a cosponsor is now one of many cosponsors and because there is an increasing likelihood that they may be notified of the bill by someone other than the original sponsor. The weighted relationship

$$A_{ij} = \sum_{b=1}^{B} \frac{w_{ijb}}{c_b},$$ (1.17)

gives the ij^{th} element of an adjacency matrix A by summing w_{ijb}, an indicator if representative i cosponsored j's bill for each bill b (note that the indicator is zero if j did not sponsor the bill in question). Each bill is weighted by the number of cosponsors c_b on the bill. This metric implies close connections between legislators i and j if they cosponsor each other's bills frequently and if there are few other cosponsors on these bills, and a weak relationship if they rarely (if ever) cosponsor each other's bills and when they do there are many cosponsors.

Cosponsorship networks are built on the idea that cosponsoring a bill is a clear signal of a relationship between the cosponsor and the sponsor. Since many bills are introduced

---

[28]This is one way to ensure that the bill makes it to the floor and will pass. Additionally, legislation that is widely popular with the public often has many cosponsors, which allows legislators to claim credit when up for reelection.

in a legislative session, repeated cosponsorship is a strong signal of a connection from the cosponsor to the bill's sponsor.[29] Fowler (2006a) describes several key attributes of cosponsorship networks; they tend to be denser and more closely connected than other social networks that are commonly observed.[30] While a cosponsorship relationship is directed, Fowler also notes a strong tendency for reciprocal cosponsorship, meaning that A_{ij} and A_{ji} tend to be correlated though not exactly equal to one another.

The cosponsorship networks from the 93rd to the 110th Congresses are combined with roll-call voting data from the same time period. The outcome variable of interest is voting yes on a bill; this outcome is aggregated over the entire congressional session (or smaller interval of time if desired) so that the final dependent variable is the total number of yea votes for each legislator in each period. Initially, I focus on a single covariate, an individual's ideology as measured by Poole and Rosenthal (2000). The exact variable used is the number of predicted yea votes, i.e. how would an individual with this ideology vote on the issue at hand?[31] This measure has several key benefits. First, it varies across every bill, which makes for a stronger relationship than most other

---

[29]The exact nature of the relationship is unclear, it may be a set of common ideas, a strategic political choice, or reveal an actual friendship. For the purposes of this application, the exact meaning of the relationship is unimportant. It is the fact that the relationships exist and vary in intensity that is crucial to identify peer effects based on this social network.

[30]More closely connected in this context indicates that these networks have a small diameter.

[31]The Poole and Rosenthal system is based on the idea of ideal points, that each legislator has an ideal point and prefers legislation that lies closest to this point. The estimation procedure to find these predicted votes takes into account each legislators voting history and then observes how everyone else voted on the bill. Based on these votes, the model will predict a yea or a nay vote for that legislator. This model is able to predict the correct vote 90% of the time without any additional information about the representatives.

attributes that may affect voting (which do not change or change very rarely). Second, it avoids the problem of needing to know the content of the bill. Since bill content can vary widely, a yea vote means different things depending on the context. This holds even if the general topic of the bill is known, i.e. voting yea on a budget bill implies very different things depending on whether or not the bill increases or decreases spending. Third, using this variable in the regression model represents a crude rational expectations model, where legislators know how everyone else is likely to vote.[32] In this context a peer effect represents deviations from a legislator's expected behavior attributable to the choice of his or her friends to also deviate from expected behavior.

In Subsection 1.6.3, I consider some other covariates including which party controls the chamber and the representative's seniority. Being a member of the party in power is much more likely to have predictive power than simple party identification (Democrat or Republican) for explaining yea votes because the majority party is able to set the legislative agenda through their disproportionate control of committee and chamber leadership positions. This power results in significant influence over which bills make it out of committee and onto the floor for a vote.[33] Note that when the model with the individual within transformation is estimated, time-invariant individual attributes cannot be separated from this effect; consequently, party identification, race, and gender cannot

---

[32]When a roll-call vote is made, voting is public so legislators are aware of who already voted and what they chose to do.

[33]This idea is set out in Cox and McCubbins (2005), a prominent work in the study of the Congress.

be included when estimating a fixed effect model. This problem is especially salient when considering only a single Congress. While there are a few individuals who did switch parties in this time period, this did not happen frequently enough to separately identify the parameters. Additionally, in the roll-call data representatives who switch party are treated as two separate individuals making it difficult to pursue this strategy.

The details of how cosponsorship matrices are derived and how the roll call votes data were prepared are available in Section A.2.

1.6.2 A Single Congress

In many ways the cosponsorship networks combined with congressional voting provide an ideal setting to demonstrate the methods proposed here. A congressional vote is a decision where peers have strong incentives to influence one another and it is intuitive to try to model this as an endogenous peer effect. These bills are plausibly independent from one another, because bills presented one after another are seldom related. Extending this idea, it seems reasonable to assume that the set of bills voted on over a short period of time are independent from bills voted on over a different short period of time. To improve the likelihood that the network is unrelated to the unobservable attributes that explain voting this period, the cosponsorship network from the previous Congress is used.[34] Additionally, this network reflects extant social ties and

---

[34]This means that new legislators who do not have cosponsorship information from the previous period are dropped from the analysis. Retention rates in the sample are typically between 85% and 95% so

avoids the complication of observing cosponsorship ties that are part of the legislative

process that also produced the legislation up for roll-call vote.

Estimation of the basic model with a single covariate and a single congress proceeds

via Gibbs sampling as outlined in Section 1.4. In this case, the outcome is the number

of yea votes cast in a given month of the 110th legislative session. The Gibbs samplers

for both the conventional instrumental variables estimate and the Dirichlet Process are

initialized with a burn-in period of 10000 observations to minimize the impact of the

initial values and run for 100,000 repetitions, keeping every 10th sample to minimize

autocorrelation in the posterior sample. Diffuse priors are chosen for all of the parameters

estimated in the model, all regression coefficients are assumed to be mean 0 with variance

1000. The Inverse-Wishart prior is also given diffuse prior values. The Dirichlet Process

priors are left as explained in Conley et al. (2008). Additional information about the

implementation of the algorithms is available in the online appendix.

The results for the simple model are contained in Table 1.4, with each column group

accounting for an additional fixed effect.[35] The effect of interest ("Vote Peer Effect") is

estimated at approximately 0.3 once the fixed effects are accounted for, indicating that if

the legislator's friends mean number of yea votes increases by 1 that legislator's number

of yea votes is likely to increase by 0.3 holding everything else constant. The effect of

---

relatively little data is lost based on this decision. However, this does mean that Freshman legislators, who may be particularly susceptible to social pressure, are never included in the sample.

[35]Note that all of these models were estimated using both the local and global within transformation to account for contextual effects. The differences between them are minimal, the global transformation is presented here.

Table 1.4: 110th Congress: Voting Peer Effects

	Simple Model		Individual FE		Individual/Bill FE	
	Coef	SD	Coef	SD	Coef	SD
Gibbs Sample						
Vote Peer Effect	0.516	0.078	0.550	0.073	0.388	0.092
Predicted Vote	0.881	0.009	0.864	0.010	0.867	0.009
Predicted Vote Peer Effect	−0.386	0.077	−0.402	0.073	−0.258	0.084
Dirichlet Process						
Vote Peer Effect	0.501	0.081	0.563	0.077	0.298	0.101
Predicted Vote	0.904	0.010	0.877	0.011	0.879	0.012
Predicted Vote Peer Effect	−0.397	0.080	−0.425	0.078	−0.239	0.092

the predicted vote is approximately 0.9, which is consistent with it being an accurate

predictor of representatives votes. The exogenous peer effect in this case is the effect of

the an representative's peer's predicted votes on an representative's votes. The negative

coefficient in this case indicates that increasing the number of predicted "yea" votes of a

legislator's friends leads to a decrease in the number of "yea" votes the legislators cast.

This somewhat counter-intuitive result may be explained if the regression is thought

of as explaining deviations from expected behavior; a change in the observed votes of

an individual's friends (the endogenous peer effect) must be offset by a negative effect

of the number of predicted votes (the exgoenous peer effect).[36] In addition, Figure 1.1

contains the vote peer effect estimate for each of the 94th-110th Senates. Note that there

are statistically significant positive peer effects in most Senates, but the magnitude of

the peer effect varies widely over time. The causes of these fluctuations is an interesting

question that is left for future research.

1.6.3 Multiple Congresses

While it may seem implausible that cosponsorship networks are independent across

congressional sessions, this assumption may be more valid than initially suspected. First,

the assumption is not that the networks are completely independent, rather it is that,

conditional on the explanatory variables used in the model, links appear seemingly

---

[36]This issue is avoided in subsequent results over multiple networks where other variables are available
to serve as an isntrument.

Figure 1.1: Peer Effects in the 94th-110th Senates
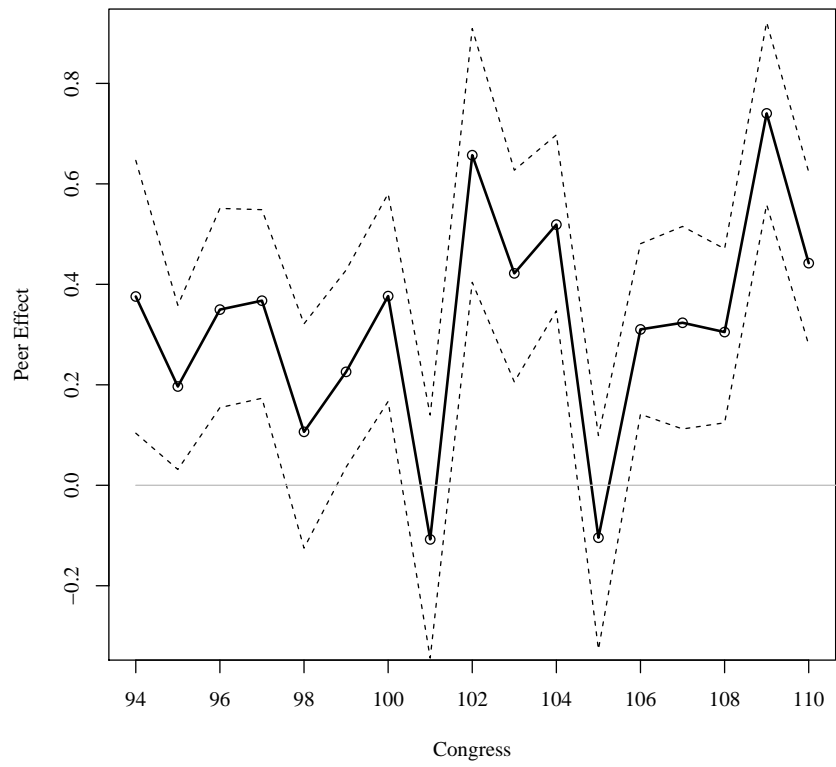
at random. This places much more weight on correctly specifying the model, but it is certainly possible to include as many individual attributes as we think might affect network formation. Second, elections between sessions can be treated as quasi-exogenous shocks to House and Senate membership. Changing the memberships will necessarily affect the structure of the network, indeed Fowler (2006a) shows that there are significant differences in cosponsorship networks across legislative sessions.

The results using multiple Congresses and a single covariate are reported in Table 1.5.[37] The outcome in these models is the sum of yea votes cast in a congressional session. In this case the model is run with two different instruments. The first (the "weak" instrument) is an indicator variable for the party that is in power. For reasons discussed earlier, this is expected to increase the number of yea votes in a given legislative session and a representative votes yea far more often as a member of the party in power; however, it also does not vary within a Congressional session and therefore is uninformative in that setting. The second instrument (the "strong" instrument) is the number of bills voted on that are sponsored by a member of your party. This variable conveys similar information as the party in power (presumably the party in power is able to bring more of their bills to a vote), but it also contains additional information about how well the party in power was able to impose its will. While both instruments should yield approximately the same parameter on the endogenous peer effect, the weak instrument in the classic

---

[37]These models account for both individual and congress level fixed effects.

Table 1.5: 94th-110th Congress: Voting Peer Effects

	Gibbs Sample				Dirichlet Process			
	Coef	SD	lower	upper	Coef	SD	lower	upper
Weak Instrument								
Vote Peer Effect	0.190	0.123	−0.076	0.416	0.159	0.093	−0.017	0.333
Party In Power	19.673	2.313	15.276	24.500	22.186	2.090	18.369	26.324
Party In Power Peer Effect	8.883	6.658	−5.807	20.158	19.273	7.571	4.850	34.356
Strong Instrument								
Vote Peer Effect	0.164	0.054	0.058	0.273	0.087	0.054	−0.030	0.183
Sponsor in Party	0.337	0.010	0.316	0.357	0.338	0.009	0.319	0.354
Sponsor in Party Peer Effect	−0.142	0.044	−0.227	−0.058	−0.062	0.042	−0.143	0.022

Gibbs sampling is unable to precisely estimate the coefficient. The Dirichlet Process appears to perform better, with the 95% Highest Probability Density (HPD) interval only just containing 0 and returning a coefficient that is closer to the estimates for the stronger instrument case.

I next estimate the model across congresses with multiple exogenous variables. Note that in this case, each variable x is included as an explanatory variable, an exogenous peer effect Gx, and an instrument G^2x which means that the model is over-identified.[38] The new variables are time in office and the ideology score. Neither variable substantially impacts the variables of interest. Time in office, measured in weeks, is significant both on its own and as a peer effect, where legislators who have been in office longer or have friends that have been in office longer are less likely to vote yes. Ideology score is insignificant, both as a covariate and a peer effect, likely because its effect varies depending on which party is in control of the chamber.

Using the methods derived in the paper, I am able to show that there are significant peer effects in congressional voting. These effects are observable within a single congress, where the network is given as the co-sponsorship relationships from the previous congressional session, and across legislative sessions. While the magnitude of the peer effect varies somewhat across specifications, the sign is consistent and the

---

[38]It is not necessary to include an explanatory variable in all three roles; however, there are reasons to prefer to do so. If the variable belongs in the model as its own effect, it is likely to have an effect through peers as well. This is especially important in variables such as income or race, where a high degree of social sorting occurs. The exogenous peer effects in these cases explicitly accounts for selection based on these characteristics in the model.

Table 1.6: 94th–110th Congress: Voting Peer Effects, Additional Covariates

	Gibbs Sample				Dirichlet Process			
	Coef	SD	lower	upper	Coef	SD	lower	upper
Vote Peer Effect	0.174	0.054	0.065	0.275	0.197	0.055	0.089	0.303
Sponsor in Party	0.452	0.012	0.430	0.476	0.455	0.013	0.430	0.481
Sponsor in Party Peer Effect	−0.160	0.051	−0.259	−0.059	−0.226	0.056	−0.333	−0.122
Party in Power	−32.334	2.036	−36.226	−28.703	−29.136	2.003	−32.999	−25.317
Party in Power Peer Effect	18.676	5.627	7.610	29.865	27.953	6.263	16.144	40.499
Time in Office	−0.008	0.002	−0.012	−0.003	−0.004	0.002	−0.008	0.000
Time in Office Peer Effect	−0.023	0.006	−0.035	−0.010	−0.021	0.006	−0.034	−0.011
Ideology Score	2.295	8.265	−13.657	18.178	−4.211	12.091	−27.335	19.699
Ideology Score Peer Effect	−10.747	6.619	−23.581	2.473	−13.232	8.163	−29.046	1.811

size of the effect is plausible. Methodologically, both methods yield results that are consistent with each other, lending credence to the idea that both estimation strategies are identifying the same effect.

1.7 Conclusion

I have presented two methods for dealing with individual heterogeneity in peer effects estimation with endogenous groups. The first, using panel methods applied to multiple networks, is able to control for individual and time period specific fixed effects and scales well to large networks; however, it requires repeated observations in networks and strong assumptions about the network formation process. The Semi-Parametric method can be implemented in a single network and is potentially able to account for individual heterogeneity that is not time-invariant, but is computationally intensive and does not scale well to large networks. Both estimation procedures work well in many settings, and often work best if fixed effects are accounted for and the model is estimated using the semi-parametric technique. I demonstrate these properties using simulated networks and then use both models to estimate peer effects in Congressional voting.

These methods can be useful in any setting where peer effects are to be estimated, especially when the peer groups are potentially endogenous. For example, online social networks fit in this framework of repeated observations of a network along with

outcomes of interest. Peer effects may also arise in purchasing decisions: how much more likely are friends to purchase an item given that that their friend mentioned buying it? Alternatively, consider a high school with a given social structure. Observing daily attendance would allow estimation of peer effects in school truancy, i.e. how much more likely is a student to miss school if their friends miss school?

Two avenues for further development of the peer effects model introduced in this work may prove especially fruitful. The first is to relax the independence across outcomes assumption to allow for time series effects when there are repeated outcomes on a fixed network. Similarly, it may be useful to merge the ideas of Section 1.2 and Section 1.3, so that we can look at multiple outcomes across multiple networks. The second avenue for future research is to integrate insights from the literature regarding longitudinal networks. Models of longitudinal network formation such as longitudinal ERGMs or SIENA[39] assume that networks continuously evolve over time and are only observed at discrete points. Such models are atheoretical, in that they do not assume any specific network formation process, but provide a basic statistical model of how networks may evolve. Combining these models with the peer effect identification procedure may provide insights into situations where a network is allowed to vary continuously.

Peer effects estimation is becoming increasingly important as technology increases the availability of network data. While much remains to be done, this paper has provided

---

[39]See `http://www.stats.ox.ac.uk/~snijders/siena/siena_r.htm` for information about SIENA modeling.

two practical empirical frameworks for estimating peer effects when panel data from social networks is available. By explicitly modeling individual level heterogeneity as part of peer effects estimation, I have been able to loosen the strict exogeneity assumption of network formation and demonstrate the usefulness of these techniques on simulated and observed network data.

Chapter 2

Social Networks and Peer Effects in Microfinance Participation

2.1   Introduction

Social influences, commonly referred to as peer effects, potentially have a large impact on individual decision making. One area in which peer effects have recently begun to receive attention is in micro-lending, especially because many microfinance progams have begun exploring group lending programs. The Diffusion of Microfinance (DoM) project examines take up of microfinance loans by looking at how information diffused among networks in villages in Southern India. DoM focuses on an experimental intervention where only certain individuals are told about the microfinance program, which allows for modeling the take-up of these individuals and their social contacts. I exploit this variation in information, along with data about social networks sampled from the villages, to estimate the effect of an individual's friends' decisions to participate in microfinance on that individual's take up of microfinance. I show, using novel network-based estimation strategies, that there are strong peer effects in microfinance participation: an additional friend choosing to participate in the program increases the likelihood that a household participates in the program by 6%-10%. These results persist across a variety of robustness checks, in which careful attention is paid to differentiating the peer effect in take-up from sorting based on observable attributes.

In addition to the substantive contribution of estimating a peer effect in microfinance participation, this work demonstrates the advantages of network based estimation

strategies in the identification of peer effects. Network approaches resolve the reflection problem and identification of the peer effect holds even when friendship formation is allowed to be endogenous, as long as the model is correctly specified. These critical specification assumptions are explored in detail, including applying recent advances in peer effects estimation that can account for unobserved individual heterogeneity. I also demonstrate techniques for estimating unbiased peer effects when the network data is incomplete. Chandrasekhar and Lewis (2012) shows that using sampled network data can results in significant biases in peer effect estimates; however, they also establish that these biases can be eliminated or minimized through analytic corrections and graphical reconstruction. Additionally, graphical reconstruction allows the inclusion of statistics about network that would be biased in a sampled network. This allows me to use network centrality to control for the probability that an individual is informed about the program.

These results help to resolve a discrepancy between the microfinance literature, which tends to show that there are significant positive peer effects in repayment decisions (Li et al., 2013; Breza, n.d.; Giné and Karlan, 2014), with the results of Chandrasekhar and Lewis (2012) which finds an insignificant or possibly negative peer effect on microfinance take-up using the DoM data. Using a more detailed specification that Chandrasekhar and Lewis (2012), I show that including a variety of individual and peer group covariates leads to estimation of positive and statistically significant peer effects in take-up decisions, consistent with prior findings regarding repayment.

I introduce the Diffusion of Microfinance project in Section 2.2 and review some recent findings about peer influences in microfinance. Section 2.3 gives background on social networks and outlines how they can be use to estimate peer effects. The primary results are contained in Section 2.4, along with the results for several alternative explanations. Section 2.5 introduces the idea of graphical reconstruction to address sampling error in networks and provides upper and lower bounds for the estimates given a variety of configurations for the missing links. Section 2.6 concludes.

2.2 Background

The focus of this work is to estimate a peer effect on participation in a microfinance program. In other words, how much more likely is an individual to take up a loan if their friends also take up a loan? Not only does this provide insight into the effects of friendship on individual behavior, but it also helps microfinance institutions better understand the dynamics of their interventions.

2.2.1 Diffusion of Microfinance

Previous work in this area has shown that social networks are very important to the implementation of microfinance lending programs. The Diffusion of Microfinance project (DoM) is a randomized experiment that attempts to determine how information

about a new lending program spreads through a village.[1] Prior to entering the village,

DoM surveys approximately half of the villagers about their social interactions and

other demographic characteristics. The microfinance institution then enters a village

by contacting a few prominent individuals (referred to as leaders) and inviting them to

inform their friends about the new program. Although the method of entering a village

was common across all the villages, the actual take up of the program varied widely,

with approximately 5% of households participating in one village with participation

nearing 50% in another.

A few research findings regarding take up of microfinance have already been estab-

lished using this data.[2] Banerjee et al. (2013) shows that the primary determinant of

program take-up rates is how prominent the informed individuals are in the village social

structure, as measured by a concept called centrality, and examines the performance

of various measures of centrality in predicting diffusion. They also fit a contagion

model that explains how information is passed from person to person. Chandrasekhar

and Lewis (2012) uses the DoM data to make several methodological contributions

(including results regarding peer effects estimators) about using network data that are

derived from a sampled network.

---

[1]The data for this project has been made available online, see Banerjee et al. (2014).

[2]These villages and the DoM data have been used in other settings as well, for an empirical example of
a favor trading model (Jackson et al., 2012), as the site of various experiments including some regarding
social insurance (Chandrasekhar et al., 2011a,b) and investment decisions (Breza et al., 2013).

While I also use the DoM data, there are several substantive differences in the analysis below. First, the primary concern of Banerjee et al. (2013) is a village level outcome: how does information about the program (proxied by takeup decisions) diffuse through the village? Whereas I am interested in the individual level, isolating the effect of friendship on takeup decisions. Second, the focus of Chandrasekhar and Lewis (2012) is methodological and the peer effect estimation is primarily to demonstrate the feasibility of their techniques. Relatively little attention is paid to the specification of the model, even though this is critical to the identification assumptions for estimating peer effects. Most importantly, the peer effect models in that work contain a single covariate which is insufficient to fully account for all of the individual attributes that might affect microfinance take-up. One reason that they may focus on a single covariate is that this variable, "whether or not the household was informed" about microfinance is assigned by the experimenter. While random assignment can often allow consistent estimation of peer effects when peer groups are also randomly assigned, this is not the case when network ties are determined endogenously. What may appear to be a peer effect may actually be due to factors that are unaccounted for in the model that affect who individuals choose as friends.[3] Consequently, the identified effect was negative and usually statistically insignificant. In contrast, I find significant positive peer effects

---

[3]This idea is discussed in more detail in Section 2.4

combining more robust model specifications with several of the techniques proposed by Chandrasekhar and Lewis (2012).

2.2.2   Peer Effects in Microfinance

Peer effects in microfinance decisions are an area of substantial interest, but there are often difficulties in getting the data necessary to properly estimate them. Using networks and the unique structure of the DoM program, I am able show that there are significant social influences in areas of micro-lending that have not been studied in depth. Much of the work to date has primarily been concerned with repayment, rather than take up, of microcredit.[4] Microfinance institutions have begun investigating group, rather than individual, liability in loans. The initial results seem to show little difference in repayment rates (see Attanasio et al. (2011) and Giné and Karlan (2014)), which would seem to indicate that social pressure in this context has minimal effect. However, these results are tempered by the fact that individuals are selecting into microfinance programs based on the knowledge of the liability arrangement.

Results that examine unexpected changes in repayment probability yield much stronger estimates of peer effect estimation. Giné et al. (2011) exploits an exogenous shock on the probability of repayment under group liability by exploiting a religious declaration that all Muslims in a province in India should stop repaying microfinance

---

[4]Banerjee (2013) contains a nice review of the current state of the microfinance literature. I focus here only on those studies that explore peer effects and influences.

loans. They find that Muslim dominated groups defaulted at higher rates than Hindu dominated groups, a result at least partially attributable to social pressure. Intriguingly, Breza et al. (2013) exploits a similar policy enacted by the state of Andhra Pradesh which encouraged all its citizens to default on microfinance loans. Despite these loans being subject to individual liability (and the policy being independent of religion), Breza demonstrates that having all of one's peers in an investment group repay resulted in a 10%-15% increase in the probability that an individual repaid their own loan. Similar results are shown by Li et al. (2013) who explore peer effects in repayment to group-lending program using a structural model. They demonstrate that if all members of the group fully repaid, the probability that an individual in that group repaid increases by about 12%. Although methodlogically distinct, my estimates of peer effects in take-up are similar to these estimates of repayment.

While there has been relatively little work regarding microfinance take-up, there is a large literature indicating that peer influences are important for individuals making investment decisions.[5] This includes results from a variety of controlled experiments (Duflo and Saez, 2003; Beshears et al., 2011; Bursztyn et al., 2014) and observational studies such as (Hong et al., 2004, 2005; Brown et al., 2008). I contribute to this literature by providing an examination of peer effects specific to microlending decisions. Additionally, this is among the first studies to use networks to identify these peer effects.

---

[5]There is obviously a large literature identifying peer effects in a variety of settings, a review of which is beyond the scope of this work.

2.3 Estimating Peer Effects in Networks

2.3.1 Networks

A network consists of a set of nodes and set of links between them. In this setting, a network consists of households (nodes) in a village in India with links between them defined as whether or not individuals in the household indicate they are friends.[6] A network is often represented by an adjacency matrix (A), where A_{ij} indicates if household i and j are friends. The row-vector $\mathbf{a}_i$ indicates all of the friends of individual i. Network links can also be weighted, for example by a count of how often two individuals interact, to indicate that some links may be more important than others. Although the networks provided by DoM are undirected, meaning that the presence of a link A_{ij} implies the presence of A_{ji}, networks can also be directed.[7] In practice this means that friendships are not required to be reciprocal.

One of the primary concerns in peer effects estimation is missing data in the network. While the DoM project has a census of households (meaning that no nodes are missing from the network), the network links are sampled for only a subset of this population.[8]

---

[6]The specific definition of friendships in this setting are discussed in detail in Subsection 2.4.1

[7]It may be preferable to estimate peer effects using a directed network as it makes the conditions for identification easier to satisfy, but the data is only available as an undirected network. In working drafts of the paper, Banerjee et al. (2013) argue that while the questions eliciting the network are directional, i.e. who do you borrow rice from, the communication opportunity is bidirectional. Additionally, they contend that since many of the links are reciprocated, along with the fact that questions about relatives and the questions about friendship have roughly the same rate of reciprocity, that lack of reciprocity is likely measurement error.

[8]See Section 2.4 for a detailed description of how the data is collected.

(A)   (B)
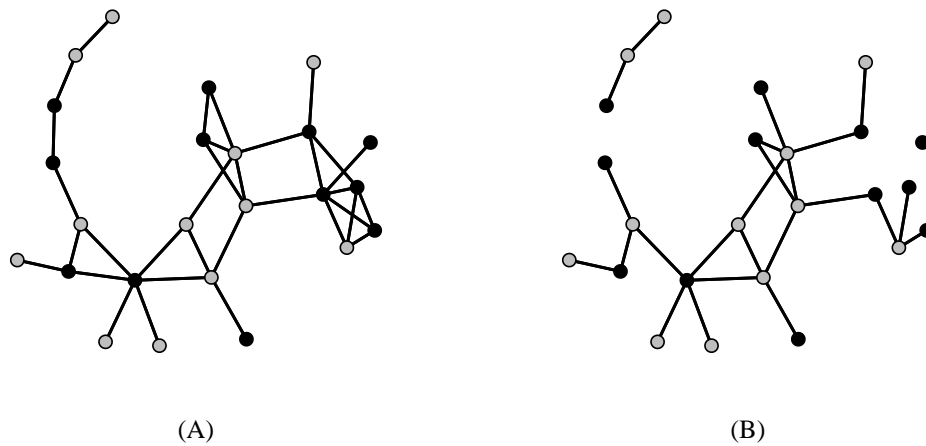
Figure 2.1: This figure shows the effect of sampling schemes on a given network. The complete network is in (A) and the sampled nodes are highlighted in gray with unsampled nodes in black. Panel (B) shows the links that are in the "star" subgraph, the network that results from only observing links nominated by sampled individuals. Observe that there are no links between black nodes, they can only link to gray nodes when the gray node nominates them as a friend.

For this subset, the network is egocentrically sampled, meaning that an individual lists all of their friends from the known set of all villagers. Figure 2.1 shows an example of an egocentric sample of the network. Panel (b) shows the complete network, with the sampled subset highlighted in gray. Panel (d) highlights the links that are observed in the sample. The key insight of Chandrasekhar and Lewis (2012) for estimating peer effects in this kind of sample is that we know all of the friends of the sampled individuals; consequently, peer effects can be estimated using the "star" sub-network shown in panel (e) as long as we only model peer effects for those whose complete friendship list is observed (the red nodes in the figure).[9] It is important to note that using this analytical correction requires that non-network data (individual covariates beyond the friendship links) be observed for all individuals in the network. Without this data it is impossible estimate exogenous peer effects properly. While this may seem like a restrictive condition, this data can be obtained in a variety of ways. For example, surveyed individuals can provide information about the friends that they have nominated. In the case of DoM, there are a set of variables observed for all households in the village which will be the primary data source for the analysis herein.

An additional method for dealing with sampled data suggested by Chandrasekhar and Lewis (2012) is graphical reconstruction. Graphical reconstruction consists of predicting whether or not links between two non-sampled individuals are likely to form. While

---

[9]In the star subgraph, the only links that are not observed are those between two individuals who are not sampled.

Chandrasekhar and Lewis (2012) do this with a simple logit model, I instead choose to simulate multiple networks that share the observed attributes of the sampled network. This allows me to bound the estimates above and below, giving guidance as to what the effect might be for a variety of possibilities for the missing links. This is discussed in detail in Section 2.5.

2.3.2 Peer Effects Model

The linear-in-means peer effects model was popularized by Manski (1993) and has subsequently been explored in a variety of contexts. This model takes the form:

$$y_{iv} = \alpha + \beta \bar{y}_{(i)} + \gamma x_i + \delta \bar{x}_{(i)} + \lambda_v + \varepsilon_{iv} \tag{2.1}$$

where an individual's outcome is a function of their peer group's mean outcome $\bar{y}_{(i)}$ ("the endogenous peer effect"), an individual observable x_i, and the peer group's mean observable $\bar{x}_{(i)}$ (the "exogenous peer effect").[10] Manski's original insight (as well as subsequent work by Moffitt et al. (2001)) is one of non-identification, i.e. that it is impossible to separately identify parameters on the endogenous and the exogenous peer effect. There are three reasons for this: the reflection problem (an individual's outcome enters on both sides of the model), contextual effects (individuals in a peer group may share unobservable shocks that can be confounded with peer effects), and

---

[10]Manski (1993) originally calculated peer effects as containing individual i; however, subsequent work (see Moffitt et al. (2001) suggests a formulation where individual i is excluded from calculating the mean. I follow the second convention here, where $\bar{y}_{(i)}$) is the mean of the peer group excluding individual i.

endogenous group membership (individuals may select peer groups because they share similar attributes). Any attempt to estimate peer effects must credibly address these three problems. Consequently, a variety of approaches have been developed to resolve them, including varying group sizes (Lee, 2007), variance-covariance restrictions (Graham, 2008), properties of binary response models (Brock and Durlauf, 2007), and social networks (Bramoullé et al., 2009; De Giorgi et al., 2010).

Bramoullé et al. (2009) shows that peer effects can be separately identified using a social network as a peer group as long as the social network has certain properties.[11] This result nests the results of Manski (1993), Moffitt et al. (2001), and Lee (2007) as special cases and shares the identification properties of those models. Additionally, as long as multiple networks are observed, it is possible to account for contextual effects using a network level fixed effect. Formally the condition for identification is that an identity matrix, the adjacency matrix, and the adjacency matrix squared must be linearly independent. In practice, this means that the network must posses intransitive triads, sets of three individuals in an undirected network where individual i is connected to individual j and j is connected to k, but i is not connected to k.[12] Additionally, a sufficient condition is that the longest "shortest' path (known as a geodesic) between any

---

[11]Although similar results were proven by De Giorgi et al. (2010), I rely on the formulation and results of Bramoullé et al. (2009) through the remainder of this paper.

[12]An intransitive triad is contrasted with a triangle, where all individuals are connected. The conditions for a directed network are slightly different, but because the networks in question are not directed, these conditions will not be explored in this work.

two individuals has length longer than two (or three in the case of the model including

fixed effects).[13] This condition is satisfied in all of the networks in question.[14]

Importantly, the Bramoullé et al. (2009) framework allows for network links to

be endogenous (formed on the basis of observable attributes); however, this requires

the assumption that the attributes that affect network formation are included in the

model. If they are not, the peer effect estimation may simply be identifying the effect

of sorting along unobserved attributes and the estimate will be biased. Prior research

using microfinance (as well as the DoM project specifically) has done little to explore

this specification problem. My primary contribution in this paper is to carefully address

how the peer effects models are specified, relying on previous results regarding social

networks to resolve the reflection problem and contextual factors. This is done by

explicitly controlling for peer observable attributes showing that effects persist across a

variety of specifications of the network. I also put forth a placebo test to show that the

peer effects identified are not simply due to endogenous sorting. Additionally, relying

on the results of Banerjee et al. (2013), I specify the model using network centrality as

a proxy for the probability of being informed. This is important to show that the peer

effect is not due simply to a higher likelihood of being informed about the program;

---

[13]The longest geodesic is calculated by identifying the shortest path between all possible combinations of nodes. If a geodesic exists that is longer than 2, there exists at least one intransitive triad.

[14]Because the networks are sampled the longest geodesic is difficult to calculate, since many links are missing which may facilitate shorter paths between individuals. However, the longest geodesic even when restricting the network to only sampled individuals is still at least 4 in all cases (it is at least 6 when using the star sub-networks). Additionally, the conditions are satisfied in the graphically reconstructed networks, which should be similar to the whole network.

however, this requires special care when using a sampled network to ensure that the network statistics are estimated properly.

Peer effects must also be estimated using a instrumental variables strategy because the outcomes appear on both sides of the model. Following work in spatial econometrics, Bramoullé et al. (2009) suggests that the adjacency matrix raised to a power and then multiplied an exogenous covariate vector is a valid instrument. This exploits the fact that the square of an adjacency matrix returns a matrix indicating the friends of an individual's friends. Since every individual's outcome in a peer effects model is a function of their friend's attributes, this is a valid instrument as long as friends of friends do not directly affect an individual. If this is the case, friends of friends only influence an individual through the individual's friends' outcomes. While the proper estimation strategy is a matter of some dispute, Lee (2007) and Bramoullé et al. (2009) prefer maxmium likelihood estimation while Goldsmith-Pinkham and Imbens (2013) suggest Bayesian models, I rely on two stage least squares (2SLS). The results vary little depending on the estimation technique; however, one reason to prefer 2SLS in this setting is computational feasibility, as the graphical reconstruction process requires estimating the model many times. Additionally, while the results should apply generally, Chandrasekhar and Lewis (2012) only prove the efficacy of their corrections for 2SLS and Generalized Method of Moments estimators.

2.4   Data and Results

A microfinance institution was considering expansion to 77 villages in the Karanataka state in India. Prior to expansion, DoM collected data from these villages including a complete household census and a detailed survey to a random sample of approximately half of the households in all of these villages. This survey collected comprehensive demographic information and asked about a variety of social interactions from which it is possible to create a social network. The microfinance institution subsequently entered a subset of the villages and documented which households had members participate in the microfinance program.

It is important to note that the social network data (and detailed demographic data) have significant limitations. The first is that they are sampled on an individual level, but the outcome is observed only on a household level. Consequently, friendship networks are aggregated to a household level by defining two households as linked if any individuals within the household report being linked to an individual in another household. The second is the missing data problem, as nonsampled individuals' friendship lists are likely missing or incomplete (because the network is undirected, if a sampled individual nominates them as a friend this relationship is contained in the data). As discussed in Subsection 2.3.1, I resolve this by focusing the analysis on the star subgraph, the sampled individuals and their complete list of friends.

Because the full set of covariates is only available for a subset of individuals in each network, I focus on the covariates available from the census of households. Summary statistics for these variables are available in Table 2.1.[15] These variables include whether or not the household contains a leader,[16] housing information (# of rooms in the house, # of beds in the house, electricity indicators and latrine indicators), household covariates (caste[17] and whether or not the household was surveyed), and eigenvector centrality.[18]. There are many observations missing caste data, including some entire villages. Caste is a strong indicator of friendship links and likely has an impact on microfinance take up; consequently, I proceed estimating models both for the entire sample (omitting caste entirely) and for the subsample where caste is completely observed in the village. This cuts the sample roughly in half.

2.4.1   Estimating Peer Effects

In this context, a network consists of a set of households with links defined between them depending on how they responded to a series of questions. These questions include

---

[15]Many variables of interest are available only in the detailed sample, including household gender composition (the microfinance institution primarily works with women), household savings, and labor market status.

[16]Leaders are defined teachers, politicians, or self-help group leaders. The microfinance group selected some of these to recruit individuals to the program.

[17]Note that religion would presumably have an effects in these models; however, over 95% of the sample is Hindu with a significant Muslim population in only a few villages.

[18]Eigenvector centrality is shown by Banerjee et. al. 2013 to best reflect how closely an individual is linked to others in the network and that it is highly correlated with being informed about the program. Since I am most interested in take up, rather than informational diffusion, I will use this as a control variable. The centrality measure is discussed in detail in Section 2.5

questions about exchange (who would you borrow kerosene from?) and social interaction

(who do you go to for advice?).[19] Following Banerjee et al (2013), I use the network that

is identified by taking the union of all the possible relationships.[20] In other words, in the

adjacency matrix $A_{ij} = 1$ if individual i answers yes to any of the possible link questions.

Additionally, recall that this matrix is undirected, $A_{ij} = A_{ji}$, meaning that all links are

assumed to be reciprocal. As outlined in Subsection 2.3.2, peer effects are identified

because the friendship networks in question satisfy the requirements of Bramoullé et al.

(2009).

2.4.2   Primary Specification

I estimate the following peer effects model:

$$y_{iv} = \alpha + \beta \sum_{j \neq i} a_i y_{jv} + \gamma x_i + \delta \sum_{j \neq i} a_i x_{jv} + \lambda_v + \varepsilon_{iv} \tag{2.2}$$

where $\mathbf{a_i}$ is a vector (the ith row the the adjacency matrix) where the jth element indicates

if individual i is friends with individual j. The coefficient β is the endogenous peer

effect, the effect of an individual's friends' choices on the individual's outcome. In

---

[19]A complete list of questions is available in Section B.1.

[20]This likely represents the most accurate representation of the true network. The fact of the matter is that specifying friendship networks is fraught with difficulty and all decisions have trade offs, i.e. assuming friendships are reciporical could mean that some links are overcounted because they are only valued by one person but omitting these links is just trading for another kind of measurement error. One indication that these effects are robust to this kind of measurement error is that they persist when using a variety of different questions to proxy for the network. I do not report these because it is possible that one specific network fits better by chance, and I do not wish to mine the data to find the network that fits "best."

this case, this is the effect of a household's friends taking a microfinance loan on the likelihood that the household takes a microfinance loan. This is contrasted with the exogenous peer effect δ, which is the effect of an individual's friends observables on the outcome.[21] For example, how much more likely is a household to take up microfinance if their friends are leaders in the village? Depending on the specification, I also include a village fixed effect λ_v. It should also be noted that I make a significant deviation from the traditional peer effects model, in that I consider the endogenous peer effect to be the sum of an individual's friends' outcomes (rather than the mean). This is primarily done to facilitate interpretation: β represents the effect of one additional friend taking up microfinance.[22]

The results of estimating equation (1) using a two stage least squares approach to account for the endogenous right hand side variable for the whole sample as well as the subsample where caste is observed are contained in Table 2.2. Column (1) replicates the results of Chandrasekhar and Lewis (2012) that find no or a potentially negative peer effect using a single covariate.[23] However, including individual and peer group

---

[21]In the analysis below, I include an exogenous peer effect for every individual covariate in the model. While this is not strictly necessary, it is good practice to include them to ensure that these effects are not confounded with the endogenous peer effect.

[22]A mean for the endogenous peer effect is easily obtained by row-normalizing the adjacency matrix (dividing a row of the matrix by the sum of the elements of the row). These results yield similar statistical significance and marginal effects of coefficients, although the magnitudes change due to the differing interpretation.

[23]The results do not match exactly, but are very similar. This is likely due to slight differences in the composition of the sample (the program has become available in additional villages since the original paper was written, but it is not apparentwhich villages these are.). Additionally, I am limited to observing leader status, which is only strongly correlated with being informed, while the Chandrasekhar and Lewis (2012) use a stronger measure ("whether or not an individual is informed") as their covariate.

covariates (exogenous peer effects) to properly specify the model indicate a large and statistically significant peer effect in take up of microfinance. One additional friend taking a microfinance loan increases the probability that an individual participates by about 7%.

The coefficients on the covariates fit with expectation. Leaders are more likely to participate in the microfinance program, which reflects their informational advantage as a subset of these individuals were targeted to receive information about the program. The loans are primarily taken up by those of lower economic status. For example those of "general" (GM) or "other backward" castes are less likely to take loans, in other words loans are more likely taken up by individuals in scheduled castes (the omitted group), those recognized by India's constitution as historically disadvantaged. Additionally, loans are more likely to be taken up by those without access to in home sanitation.

Lastly, I include village fixed effects to account for contextual factors, i.e. factors which impact all individuals in a given village. The existence of contextual factors is one of the classic threats to identification of peer effects models outlined by Manski (1993) and Moffitt et al. (2001). Since I have observations on multiple villages, I am able to account for these contextual factors. Running the same models on the subset of data where caste is observed yields broadly similar results, although the peer effect is no longer significant when village fixed effects are added (possibly due to the decreased sample size)

2.4.3 Network Specification

There are many possible ways to specify the network, and I consider many different ways to do so. I primarily use a network that has friendship indicators as binary. It is also possible to use the weighted network that treats A_{ij} as a count of the number of questions that i indicated a friendship with j. The weighted network gives additional weight to friendships that are presumably closer, but the results are similar enough that I have not reported them.

Since the microfinance program is primarily directed at women, it is possible that female friendships are more meaningful in this context. Table 2.3 reports the results for the model using only links that are nominated by female household members. It does appear that the peer effects are slightly stronger using this definition of the network.

2.4.4 Placebo Test

There remains a concern that, rather than estimating a endogenous peer effect in microfinance take-up, the model is yielding false positives based on preexisting factors that I am unable to fully account for in the model. Since the identification strategy relies on the timing and random selection of individuals to be initially informed, it would be concerning if an identical specification estimated positive peer effects for an outcome that should be independent of the timing and initial injection points. One way to check for this is to see if the model returns positive coefficient estimates for an outcome that

should be unrelated to the implementation of the program. One such variable, suggested by Banerjee et al. (2013), is whether or not a household has a slate roof. This variable is likely concentrated among wealthy households, and may even exhibit some sort of peer influence; however, an individual's choice of roofing material is determined prior to the programs inception and already known to all.[24] Consequently, any peer effect due to installation of a slate roof should already have occurred.

I run an identical specification; however, this time I use as the outcome whether or not the household has a slate roof. The results are presented in Table 2.4. While there is a peer effect estimated in models that do not account for contextual effects (columns 2 and 5), this effect is small and indistinguishable from zero in the fixed effect models. Presumably, the positive and significant coefficient is obtained due to differences across villages in the extent of tile roofing.

2.5   Graphical Reconstruction

The social networks are collected from individual surveys of a sample of households and are therefore incomplete. This can be resolved by limiting the sample as in Section 2.4; however, this precludes the use of any network statistic that relies on knowledge of the entire network. In this specific case, it may be desirable to account for the likelihood that an individual is informed about the program so that we isolate the

---

[24]The data was collected prior to the programs entry to the villages.

effect of peer decision making and not the fact that some individuals may simply have been uninformed. Banerjee et al. (2013) show that one of the best measures to account for the likelihood of being informed about the program is eigenvector centrality.

As the name would imply, eigenvector centrality is calculated from the eigenvectors of the adjacency matrix.[25] Intuitively, it can be thought of as a measure of how well connected an individual is in the network. As such, someone with high eigenvector centrality is likely to be informed about the program because they have many connections, someone with low eigenvector centrality is unlikely to be informed. Unfortunately, eigenvector centrality is a measure that depends on knowing the whole network (for example, someone may be highly central, but unsampled, and would have their centrality underestimated accordingly.) This sampling problem can be resolved using a model to obtain consistent estimates of which missing links are likely to form.

I estimate the probability a link forming between two households using an exponential random graph models (ERGMs).[26] ERGMs nest a logit model as a special case, but generalize to allow the use of network statistics in the model. Unfortunately, when the network is egocentrically sampled many network statistics that might be of interest are biased; however, Krivitsky et al. (2011) indicate that the degree distribution (the

---

[25]For details about the measure, along with other measures of centrality, the reader is directed to Jackson (2008)

[26]The critical attribute of the model is that it provides consistent estimate of the likelihood of link formation, while such results were not established prior to Chandrasekhar and Lewis (2012), they have been subsequently established in Chandrasekhar and Jackson (n.d.). The results are similar when using other methods suggested by Chandrasekhar and Lewis (2012), including an link-independent logit model.

distribution of the number of links each individual has) and mixing distributions (the number of links between individuals of different types) are unbiased and can be used to estimate the model. These models are very similar to what would be expected from a logit model explaining link formation. A mixing matrix condition might imply that the probability of a link forming is higher if the two households match on the observed characteristic, for example if they are from the same caste.

I fit an ERGM to each village that has complete information about caste and predict probabilities of link formation for all connections between unsampled individuals (connections between sampled individuals are assumed to be correctly observed.) I then estimate the model using the reconstructed network. It would be ideal if graphical reconstruction were to work for the peer effect as well; however, no such results have been established. This is because the network affects peer effect estimation in two ways. It is used to calculate the actual peer effects and has an important role in generating the instrument, which means that the first stage of the two stage least squares may have bias as well. Because of this, I provide two sets of results, those with the whole network in columns (1) and (3) and those that are again restricted to the surveyed individuals in columns (2) and (4). The first pair should be considered speculative in regards to the peer effect estimation, while the second set should yield correct peer effects estimates due to its focus on only sampled individuals.

The estimates including eigenvector centrality are provided in Table 2.5. The first two columns are the results of the model when estimated without reconstruction, the reconstructed network is used in columns (3) and (4). The estimates of peer effects are consistent across the models; however, as is expected the coefficient estimates for eigenvector centrality vary widely. They are presumably correctly estimated in model (4), where the sample is limited to only surveyed individuals (thus giving a proper estimate of peer effects) and the graphical reconstruction corrects for the bias when using eigenvector centrality in a sampled network. Column (4) indicates that individuals with higher centrality are more likely to participate in the program, which is partially due to them being more likely to be informed. Crucially, peer effects are still strongly present in this model.

ERGMs also provide a means of simulating many networks that fit the model of link formation. In addition to using the graphical reconstruction technique, I also propose simulating many random networks using an ERGM and estimating the 2SLS model for each. This is done to provide some intuition about how the estimates may vary based on potential configurations of the unobserved links. While the graphical reconstruction process establishes that the coefficient estimates will be consistent, little is known about how such estimators behave when the predictive strength of the reconstructive model varies. Simulating many possible networks can provide additional intuition in the case where the model of network formation is weak. In the ideal case, a strong model of

network formation will yield clear indicators of which links should be present and the estimates will be perfect. Since friendship formation processes are often driven by unobservables, these models rarely yield such strong predictions. Simulating many networks allows us to obtain bounds on the parameters that account for some of this randomness.

The bounded coefficients for the entire village are found in Table 2.6 while the results restricted to only sampled individuals are in Table 2.7. The results indicate that the effect of eigenvector centrality in the model varies widely, but that peer effect estimates are far more consistent in this model. Intriguingly, the peer effect estimate results for the whole sample in Table 2.6 (which are speculative) seem to behave quite well in that the results mirror quite closely the results on the surveyed individuals only that are known to be consistent (Table 2.7).[27]

2.6 Conclusion

I use new network based peer effect estimation strategies to estimate significant peer effects in microfinance take-up. An additional friend participating in micro-finance increases the probability that an individual participates by 6%-10%. The magnitude of this effect fits with previous results regarding microfinance repayment, and is robust to sampling error introduced by the data collection process. Most importantly, the results

---

[27]Formally establishing these results is an interesting avenue for further research.

indicate the importance of careful specification of peer effects models as omitting factors that may affect friendship formation biases the results.

This research may also have important implications for microfinance institutions and others wishing to design interventions that have maximum participation. The choice of a single individual to take up microfinance has significant spill-overs into the behavior of the individual's friends. Knowledge of the extent of this effect could help forecast participation rates with more precision. In addition, the results from the graphical reconstruction suggest avenues for future methodological research, for example, directly establishing the properties of graphical reconstruction for the estimation of peer effects.

Table 2.1: Summary Statistics

	All Households		Full Sample		Caste Observed	
	mean	sd	mean	sd	mean	sd
Participation						
Participation	0.175	(0.380)	0.184	(0.388)	0.173	(0.378)
Peer Participation	1.769	(2.134)	2.678	(2.492)	2.550	(2.417)
Covariates						
Leader	0.125	(0.331)	0.140	(0.347)	0.135	(0.342)
Number of Rooms	2.381	(1.320)	2.441	(1.346)	2.487	(1.259)
Number of Beds	0.855	(1.386)	0.892	(1.340)	0.767	(1.111)
No Electricity	0.0653	(0.247)	0.0589	(0.235)	0.0527	(0.223)
No Latrine	0.723	(0.447)	0.712	(0.453)	0.750	(0.433)
GM or OBC	0.650	(0.477)	0.646	(0.478)	0.679	(0.467)
Eigenvector Centrality	0.0526	(0.0457)	0.0777	(0.0487)	0.0750	(0.0481)
Exogenous Peer Effects						
Leader	1.607	(1.755)	2.536	(1.923)	2.521	(1.856)
Number of Rooms	24.85	(20.65)	38.28	(20.70)	39.74	(20.53)
Number of Beds	9.502	(10.56)	14.57	(12.10)	13.00	(10.56)
No Electricity	0.493	(0.941)	0.772	(1.157)	0.673	(1.076)
No Latrine	6.681	(5.698)	10.25	(5.885)	10.78	(5.871)
GM or OBC	6.786	(6.541)	10.44	(7.174)	10.44	(7.174)
Eigenvector Centrality	0.789	(0.692)	1.171	(0.744)	1.152	(0.746)
Observations	10087		4892		2240	

All households refers to all households in all of the villages. Surveyed households limits the sample to those surveyed to obtain friendship networks. Caste observed contains only households in villages where caste is completely observed. "GM or OBC" refers to castes that are not eligble for affirmative action; the whole sample and surveyed household column values of "GM or OBC" are the average excluding those with missing values.

Table 2.2: Peer Effects in Microfinance Participation

	Full Sample			Caste Observed		
	(1)	(2)	(3)	(4)	(5)	(6)
Endo. Peer Effect						
Peer Participation	-0.058*	0.086***	0.072**	-0.084	0.078***	0.031
	(0.035)	(0.017)	(0.037)	(0.070)	(0.011)	(0.036)
Covariates						
Leader	0.038*	0.027*	0.032**	0.067*	0.039	0.050*
	(0.021)	(0.015)	(0.015)	(0.034)	(0.026)	(0.029)
Number of Rooms		-0.002	-0.005		0.000	-0.001
		(0.005)	(0.005)		(0.006)	(0.006)
Number of Beds		-0.003	-0.002		-0.014	-0.018**
		(0.005)	(0.005)		(0.009)	(0.009)
No Electricity		0.031	0.027		0.069**	0.056**
		(0.022)	(0.021)		(0.030)	(0.027)
No Latrine		0.047***	0.035**		0.026	0.024
		(0.015)	(0.015)		(0.016)	(0.017)
GM or OBC					-0.004	-0.023
					(0.026)	(0.028)
Exogenous Peer Effects						
Leader	0.021*	-0.006	-0.000	0.034*	0.005	0.016*
	(0.011)	(0.005)	(0.008)	(0.018)	(0.007)	(0.009)
Number of Rooms		-0.003***	-0.004***		-0.004***	-0.003*
		(0.001)	(0.001)		(0.001)	(0.002)
Number of Beds		0.001*	0.001		0.002**	-0.000
		(0.001)	(0.001)		(0.001)	(0.002)
No Electricity		-0.000	0.000		0.008	0.000
		(0.006)	(0.005)		(0.005)	(0.007)
No Latrine		-0.006*	-0.003		-0.003	0.006
		(0.003)	(0.006)		(0.002)	(0.006)
GM or OBC					0.002	-0.001
					(0.002)	(0.004)
Village Fixed Effects	No	No	Yes	No	No	Yes
Mean Participation	0.184	0.184	0.184	0.173	0.173	0.173
Adjusted R-squared	.	0.060	0.087	.	0.080	0.109
N	4892	4892	4892	2240	2240	2240

The outcome is whether or not an individual participated in the microfinance program. The network used to calculate the peer effect is binary, indicating if the an individual in the households indicated a friend in another household. The coefficient on peer participation indicates the effect of having one additional friend adopt microfinance.

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Standard errors clustered at the village level in parentheses.

Table 2.3: Peer Effects in Microfinance Participation, Female Network

	Full Sample			Caste Observed		
	(1)	(2)	(3)	(4)	(5)	(6)
Endo. Peer Effect						
Peer Participation	-0.065	0.107***	0.115***	-0.051	0.071***	0.044
	(0.046)	(0.021)	(0.030)	(0.048)	(0.024)	(0.032)
Covariates						
Leader	0.031	0.025*	0.027*	0.047*	0.039*	0.047**
	(0.019)	(0.015)	(0.015)	(0.025)	(0.023)	(0.023)
Number of Rooms		-0.004	-0.006		-0.000	0.000
		(0.004)	(0.004)		(0.006)	(0.006)
Number of Beds		-0.003	-0.002		-0.010	-0.014*
		(0.005)	(0.005)		(0.008)	(0.008)
No Electricity		0.036*	0.034*		0.063**	0.060**
		(0.021)	(0.021)		(0.028)	(0.027)
No Latrine		0.036***	0.029**		0.021	0.023
		(0.013)	(0.014)		(0.015)	(0.015)
GM or OBC					-0.005	-0.013
					(0.028)	(0.029)
Exogenous Peer Effects						
Leader	0.024*	-0.007	-0.005	0.029**	0.010	0.019*
	(0.014)	(0.007)	(0.008)	(0.013)	(0.009)	(0.010)
Number of Rooms		-0.003***	-0.004***		-0.003**	-0.002
		(0.001)	(0.001)		(0.002)	(0.002)
Number of Beds		0.001	0.002		0.002	-0.001
		(0.001)	(0.002)		(0.002)	(0.003)
No Electricity		0.000	-0.002		0.004	-0.001
		(0.007)	(0.007)		(0.007)	(0.008)
No Latrine		-0.007*	-0.008		-0.002	0.005
		(0.004)	(0.006)		(0.003)	(0.006)
GM or OBC					-0.001	-0.002
					(0.003)	(0.004)
Village Fixed Effects	No	No	Yes	No	No	Yes
Mean Participation	0.184	0.184	0.184	0.167	0.167	0.167
Adjusted R-squared	.	0.067	0.052	.	0.097	0.110
N	4887	4887	4887	2391	2391	2391

The outcome is whether or not an individual participated in the microfinance program. The network used to calculate the peer effect is binary, indicating if a *female* member of the household indicated a friend in another household. The coefficient on peer participation indicates the effect of having one additional friend adopt microfinance.

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Standard errors clustered at the village level in parentheses.

Table 2.4: Placebo Test: Peer Effects in Tile Roofs

	Full Sample			Caste Observed		
	(1)	(2)	(3)	(4)	(5)	(6)
Endo. Peer Effect						
Tile Roof	0.007	0.049***	0.015	-0.035	0.057***	-0.004
	(0.011)	(0.006)	(0.017)	(0.035)	(0.007)	(0.027)
Covariates						
Leader	-0.068***	-0.025	-0.020	-0.035	-0.012	-0.005
	(0.021)	(0.018)	(0.018)	(0.049)	(0.023)	(0.023)
Number of Rooms		-0.011	-0.017***		-0.019	-0.024**
		(0.007)	(0.006)		(0.012)	(0.010)
Number of Beds		-0.016***	-0.016***		-0.007	-0.004
		(0.006)	(0.006)		(0.013)	(0.012)
No Electricity		-0.032	-0.035		-0.116**	-0.123***
		(0.035)	(0.038)		(0.048)	(0.046)
No Latrine		0.160***	0.154***		0.186***	0.174***
		(0.020)	(0.017)		(0.031)	(0.029)
GM or OBC					0.016	-0.042
					(0.048)	(0.041)
Exogenous Peer Effects						
Leader	-0.013**	-0.001	0.001	0.006	0.010*	0.008
	(0.006)	(0.004)	(0.004)	(0.023)	(0.006)	(0.006)
Number of Rooms		-0.002***	-0.002*		-0.005***	-0.002*
		(0.001)	(0.001)		(0.001)	(0.001)
Number of Beds		0.000	0.000		0.001	0.001
		(0.001)	(0.001)		(0.001)	(0.002)
No Electricity		0.003	0.008		0.020***	0.022*
		(0.005)	(0.008)		(0.005)	(0.012)
No Latrine		-0.011***	0.002		-0.014***	0.007
		(0.003)	(0.006)		(0.004)	(0.010)
GM or OBC					-0.001	-0.002
					(0.004)	(0.002)
Village Fixed Effects	No	No	Yes	No	No	Yes
Mean Participation	0.332	0.332	0.332	0.410	0.410	0.410
Adjusted R-squared	0.067	0.278	0.325	.	0.296	0.343
N	4892	4892	4892	2240	2240	2240

The outcome is whether or not an individual participated in the microfinance program.

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Standard errors clustered at the village level in parentheses.

Table 2.5: Peer Effects in Microfinance Participation, Include Eigenvector Centrality

	No Reconstruction		Graph. Reconstruction	
	(1)	(2)	(3)	(4)
Endo. Peer Effect				
Peer Participation	0.062**	0.075***	0.091***	0.065**
	(0.025)	(0.026)	(0.024)	(0.033)
Covariates				
Leader	0.070***	0.043	0.061***	0.040
	(0.019)	(0.028)	(0.018)	(0.030)
Number of Rooms	-0.006	0.000	-0.006	-0.001
	(0.005)	(0.006)	(0.005)	(0.007)
Number of Beds	-0.012	-0.016*	-0.010	-0.019**
	(0.007)	(0.009)	(0.007)	(0.009)
No Electricity	0.008	0.066**	0.010	0.065**
	(0.020)	(0.028)	(0.018)	(0.028)
No Latrine	0.042***	0.024	0.041***	0.026
	(0.016)	(0.017)	(0.015)	(0.017)
GM or OBC	-0.052***	-0.011	-0.035	-0.016
	(0.019)	(0.028)	(0.023)	(0.028)
Eigenvector Centrality	1.083	-0.185	-0.025	0.415
	(1.545)	(1.589)	(1.326)	(1.535)
Surveyed	-0.029*		0.005	
	(0.016)		(0.016)	
Exogenous Peer Effects				
Leader	0.010	0.010	0.006	0.012
	(0.007)	(0.008)	(0.006)	(0.008)
Number of Rooms	-0.004***	-0.003**	-0.003***	-0.002*
	(0.001)	(0.001)	(0.001)	(0.001)
Number of Beds	0.002**	0.001	0.001	-0.000
	(0.001)	(0.002)	(0.001)	(0.002)
No Electricity	0.003	0.003	-0.002	0.001
	(0.006)	(0.006)	(0.006)	(0.006)
No Latrine	0.002	-0.001	-0.002	0.001
	(0.004)	(0.004)	(0.005)	(0.004)
GM or OBC	0.002	0.002	0.004	0.002
	(0.002)	(0.002)	(0.002)	(0.003)
Eigenvector Centrality	-0.093	-0.032	-0.066	-0.083
	(0.088)	(0.104)	(0.068)	(0.095)
Surveyed	0.000		0.002	
	(0.003)		(0.004)	
Village Fixed Effects	Yes	Yes	Yes	Yes
Mean Participation	0.159	0.173	0.153	0.168
Adjusted R-squared	0.098	0.085	0.064	0.090
N	4571	2240	4600	2146

The outcome is whether or not an individual participated in the microfinance program. These models contain eigenvector centrality, a proxy for whether or not an individual is informed about the program, which is biased in a sampled network. Sample sizes change because isolated invidiuals are excluded from the previous models, but included in reconstruction. Columns (1) and (3) contain all observations with caste observed, even those not sampled, while Columns (2) and (4) retain the estimation sample of surveyed households only.

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Standard errors clustered at the village level in parentheses.

Table 2.6: Graphical Reconstruction: Bounds on Coefficients, Full Sample

	min	p25	p50	p75	max
Participation					
Peer Participation	0.0455	0.0806	0.0880	0.0949	0.122
Covariates					
Leader	0.0510	0.0603	0.0628	0.0650	0.0734
Number of Rooms	-0.00899	-0.00679	-0.00608	-0.00546	-0.00277
Number of Beds	-0.0133	-0.0104	-0.00953	-0.00881	-0.00587
No Electricity	-0.00744	0.00393	0.00702	0.0102	0.0203
No Latrine	0.0318	0.0411	0.0429	0.0450	0.0522
GM or OBC	-0.0525	-0.0389	-0.0350	-0.0309	-0.0142
Surveyed	-0.0119	-0.00200	0.000814	0.00341	0.0124
Eigenvector Centrality	-0.869	-0.0420	0.166	0.366	1.086
Exogenous Peer Effects					
Leader	-0.00527	0.000628	0.00231	0.00390	0.0102
Number of Rooms	-0.00426	-0.00288	-0.00255	-0.00225	-0.00113
Number of Beds	-0.00112	0.000364	0.000818	0.00124	0.00300
No Electricity	-0.0102	-0.00396	-0.00174	0.000447	0.0110
No Latrine	-0.0101	-0.00481	-0.00354	-0.00218	0.00296
GM or OBC	-0.00230	0.00218	0.00304	0.00400	0.00795
Surveyed	0.000179	0.00296	0.00381	0.00466	0.00843
Eigenvector Centrality	-0.129	-0.0882	-0.0781	-0.0677	-0.0373

These are the bounds on the estimation results from 1000 networks simulated via graphical reconstruction. This table should be compared to models (1) and (3) of Table 2.5.

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Standard errors clustered at the village level in parentheses.

Table 2.7: Graphical Reconstruction: Bounds on Coefficients

	min	p25	p50	p75	max
Participation					
Peer Participation	0.0525	0.0719	0.0759	0.0802	0.102
Covariates					
Leader	0.0341	0.0376	0.0383	0.0392	0.0430
Number of Rooms	-0.00143	-0.000960	-0.000866	-0.000775	-0.000197
Number of Beds	-0.0190	-0.0183	-0.0182	-0.0180	-0.0174
No Electricity	0.0624	0.0663	0.0672	0.0680	0.0718
No Latrine	0.0251	0.0256	0.0257	0.0258	0.0265
GM or OBC	-0.0190	-0.0143	-0.0134	-0.0124	-0.00778
Eigenvector Centrality	-0.381	0.0554	0.139	0.224	0.532
Exogenous Peer Effects					
Leader	0.00754	0.00995	0.0105	0.0110	0.0136
Number of Rooms	-0.00198	-0.00187	-0.00184	-0.00182	-0.00172
Number of Beds	-0.000281	-0.0000174	0.0000413	0.0001000	0.000434
No Electricity	0.00134	0.00185	0.00196	0.00209	0.00247
No Latrine	-0.00304	-0.000424	0.000135	0.000678	0.00334
GM or OBC	0.00163	0.00301	0.00330	0.00360	0.00571
Eigenvector Centrality	-0.114	-0.0868	-0.0815	-0.0763	-0.0466

These are the bounds on the estimation results from 1000 networks simulated via graphical reconstruction. This table should be compared to models (2) and (4) of Table 2.5.

Chapter 3

Juvenile Crime and the Four-Day School Week in Colorado [1]

---

[1] **Joint Work with Stefanie Fischer**

3.1   Introduction

It is the belief of many parents, administrators, and public officials that students who spend more time in school are less likely to get into trouble with the law. This belief is based on a theory known as incapacitation, whereby students who might otherwise engage in criminal activity are deprived of opportunities to break the law. Understanding the impact of time spent in school on criminal behavior has important policy implications, especially as school districts experiment with alternative schedules (year-round school, four-day weeks, etc.) in an attempt to cut costs and/or boost student performance. In order to estimate the casual relationship between time spent in school and youth crime, variation in the school schedule is necessary. Existing studies rely on sporadically occurring disruptions to school schedules which reduce the number of days a student is in school in a given week, e.g. teacher inservice days (Jacob and Lefgren, 2003), randomly occurring events such as strikes (Luallen, 2006), or furloughs (Akee et al., 2013). We offer an alternative identification strategy, one which exploits the adoption of four-day school weeks across counties and years within the state of Colorado, one of the states where the four-day week is most pervasive.

Using variation in the number of days a student is in school each week, we estimate the effect of school attendance on juvenile crime outcomes and demonstrate that juvenile arrests for larceny increase in counties that have adopted a four-day school week. The

policy does not seem to impact other kinds of crime, as estimated coefficients for drug and violent crimes do not attain statistical significance. We also show, using data on reported crimes by day of the week, that there is not evidence of a displacement of crimes; that is, absolute crime appears to be increasing as a result of the policy, it is not just shifted from a school day to a day off.

Our study contributes to the juvenile crime and incapacitation literature in several ways. First, we study long term patterns in behavior associated with incapacitation rather than the short term responses to unexpected changes examined in prior research.[2] Presumably, parents, schools, and communities can more adequately prepare for these more permanent changes in the time students are away from school. Second, because four-day school weeks are implemented much more frequently in rural areas, we analyze effects on juvenile behavior in a setting which remains understudied relative to policies in urban areas. Finally, we show that four-day school week policies may have unintended consequences that should be considered when these policies are proposed.

The paper proceeds as follows. We review four-day school week policies and the effects of school attendance on crime in Section 3.2. We outline our empirical methodology in Section 3.3 and present the baseline results in Subsection 3.3.2. Section 3.4 shows evidence that the findings persist across a variety of robustness checks. Section 3.5 concludes.

---

[2]While inservice days (Jacob and Lefgren, 2003) are planned ahead of time by school officials, they appear sporadically throughout the year and often vary widely from location to locations.

3.2   Juvenile Crime and Four-Day School Week Policies

3.2.1   How Might the Four-Day School Week Impact Youth Crime?

There are at least two channels in which students' school attendance may affect youth criminal behavior. *Incapacitation* refers to the notion that juveniles who are kept busy will remain out of trouble; or, put differently, students who have less idle time are less likely to engage in illegal activities. School, after-school programs, and other youth programs are forms of juvenile incapacitation. In fact, it is a common perception that increasing funds to expand these types of youth programs or lengthening the time students are in school will reduce youth crime. However, the exact effect of increasing school or youth program participation remains unclear. Incapacitation is expected to reduce juvenile crime by providing students with adult supervision so they have less opportunity to commit a crime. An alternative story is that school or other youth programs increase the *concentration* of youth, sometimes bringing together diverse groups of students, which may actually promote coordination of crime or provide a space for conflict to arise.

One method of answering these questions is through experimental interventions in after-school programs; however, their insight is limited because the programs typically cannot be made mandatory and those most at risk avoid them (see Cross et al. (2009)). Empirically, incapacitation effects have been widely studied and show decreases in

criminal activity and other risky behavior such as teen pregnancy; however, the primary focus of this literature has been on variation on the number of years spent in school based on compulsory schooling laws (Lochner and Moretti, 2004; Black et al., 2008; Berthelon and Kruger, 2011; Anderson, 2012; Anderson et al., 2013). This means that the effects are often temporally distinct from the actual program occupying the individuals, so these studies are able to provide little insight into the effect of a program to keep students out of trouble in the immediate future. Additionally, they are unable to estimate concentration effects.

Research that addresses the distinction between incapacitation and concentration effects typically relies on exogenous variation in the amount of time that a student spends in school. Jacob and Lefgren (2003) rely on teacher in-service days to estimate a causal relationship between school attendance and crime. They find that juvenile property crime declines by 14% on days when school is in session but violent crime for this same group increases by 28% percent on school days. A reduction in property crime associated with increased school attendance is consistent with incapacitation. The spike in violent crime on these days provides evidence for a concentration story.

Luallen (2006) exploits school attendance variation caused by teacher strikes which resulted in canceled school days to estimate an incapacitation effect. He finds that juvenile property crime and violent crime increase on days with strikes, but that the results are solely driven by urban areas. Akee et al. (2013) estimate the school-crime

relationship based on public school teacher furlough days in Hawaii and find that time off from school is associated with significantly fewer juvenile crimes which supports a concentration story.

Using four-day school weeks allows us to make several contributions to the incapacitation literature. The first is that the intervention we use is concentrated in rural areas, meaning that we can gain insight into areas where effects on juvenile criminal behavior have been less studied. Additionally, we are able to look both at overall crime rates in counties with four-day school weeks over time, but can also examine concentration effects by looking at crime on specific days of the week.

3.2.2 Four Day School Weeks and Colorado

As of 2008, seventeen states have switched a portion of their school schedules from a five day week to a four-day week.[3] The primary motivation for states to implement this policy is to reduce transportation costs, which are especially salient to rural schools. The four-day school week became particularly popular during the energy crisis in the 1970s, at which time many states began changing laws regarding days spent in school. During this same time the Colorado legislature changed their law from a mandatory number

---

[3]Starting with South Dakota in the 1930s, the following states have schools on the four-day week: Arizona, California, Colorado, Idaho, Kansas, Kentucky, Louisiana, Michigan, Minnesota, Montana, New Mexico, Oregon, South Dakota, Texas, Utah, Wisconsin and Wyoming. However, many of these programs are not very extensive. See http://www.ncsl.org/research/education/school-calendar-four-day-school-week-overview.aspx for background on specific state legislation regarding found day schools.

of school days to a mandatory number of hours, enabling districts in the state to adopt a four-day school week. Subsequent to the law's passage, alternative schedules have increased in popularity to the point where approximately 20% of students in Colorado attend four-day schools. This study exclusively focuses on the four-day school week policy in Colorado. In Colorado, the schools who have adopted a four-day school week most often cite financial savings as their reason (Grau and Shaughnessy, 1987; Donis-Keller and Silvernail, 2009; Anderson and Walker, 2012). Other reasons schools have decided to switch include parent support, improved attendance, and increased academic performance.

Given that cost considerations are central to the decision to switch, research into four-day school weeks has primarily focused on financial savings. Grau and Shaughnessy (1987), using data from ten school districts in New Mexico, document that districts operating on a four-day week experience a 10%-25% savings on fuel, electricity and transportation costs. Griffith (2011) examines six school districts that are either on the four-day week or in transition to that schedule and finds that the policy yields a maximum of about 5.5% savings.[4] Despite their growing prevalence, little work has been done to understand the impact of four-day school weeks on students. To our knowledge, the only study at this point which evaluates the impact of four-day school weeks is

---

[4]Four day school weeks have been of interest in popular media as well and journalists have gone to some effort to examine specific cases of the policy change. A TIME Magazine article (Kingsbury, 2008) reports that some rural school districts experienced large savings on transportation, utility, and insurance costs as a result of the policy and a Wall Street Journal article (Herring, 2010) sheds light on the savings that the policy has brought to a rural district in Georgia.

Anderson and Walker (2012). They find a modest, but statistically significant, positive relationship between the policy and elementary school students' math and reading test scores. Their findings suggest that switching to a four-day week does not compromise student achievement, and may even improve it.

It is important to note that even under a four-day school schedule the number of hours a student spends at school per year remains constant and is typically set by state statute. To compensate for one fewer day of instruction, those on the four-day week schedule attend school for more hours per day and/or more days in the year. Of the schools that have switched in Colorado, roughly 80% are on a Monday through Thursday schedule with Fridays off with the remainder on a Tuesday-Friday schedule with Mondays off. In the case of an incapacitation story, theory suggests that juvenile crime will likely increase in the counties that have adopted a four-day school week schedule, specifically on the days of the week that students are not attending school (Monday or Friday) due to lack of adult supervision. Not only are students not at school, their parents are also likely at work and less likely to be aware of the students' whereabouts. On the other hand, a concentration story would imply that crime would go down in these areas because there are fewer opportunities to interact with other students. Additionally, any effects due to concentration would be observed on days that students are in school. Previewing results in Subsection 3.3.2, the increases in juvenile property crime (especially larceny) would

imply that the incapacitation effect dominates.[5] There is little evidence of any specific day of the week effects.

3.3 Data and Results

3.3.1 Data

We combine several data sources for our analysis which includes 64 counties in Colorado for the years 1993-2009. Data indicating which schools are on a four-day school week and the timing of when they switch from a five day schedule to a four-day schedule come from the Colorado Department of Education.[6] Figure 3.1 shows the changes across counties and years in the percent of students on a four-day school week in the state. Nearly half of all counties have at least one school that is on a four-day school week. Nineteen percent of students in Colorado attend a school on this schedule. This data source also links schools with counties.

We use the CCD (Common Core of Data from the National Center for Education Statistics), which contains the universe of Colorado schools, to obtain total student population by county and several measures of student body composition. The CCD and

---

[5]Jacob and Lefgren (2003) see concentration effects primarily with increases in juvenile violent crime when school is in session. We see no discernible effect on violent crime, another indication that concentration effects due to four-day school weeks are minimal, possibly because hours spent with the same group of peers remain constant in a year.

[6]We would like to thank Mark Anderson of Montana State University for helping us obtain this data.

the list of four-day schools are combined to obtain the treatment variable: the percentage

of students between grades 6 and 12 on a four-day school week. Juveniles of these ages

(11-17) account for over 99% of juvenile crimes reported in Colorado. While the results

are robust to alternate specifications,[7] we believe the appropriate treatment group for

examining the effect of four-day schools is to examine its direct effect on the population

responsible for perpetrating most juvenile crime.

We use two sources of crime data. The first is county level arrest data from the

Colorado Department of Public Safety which contains reported arrests by crime type

by county in a given year. This data is available for the duration of the sample period

and summary statistics are contained in Table 3.1. This data has the advantage of being

comprehensive, covering all arrests in Colorado; however, since it is aggregated yearly

the timing does not match up with the treatment variable which is reported by the

academic calendar. The second dataset is obtained from the National Incident Based

Reporting System (NIBRS) program. NIBRS provides detailed information on reported

crime at an incident (reported crime) level, including the date of the crime. This allows

us to match the reported crime with the treatment variable by academic year and avoid

the timing problem. We limit our analysis to the set of observations with offender

information so that we may identify the age of the perpetrator of the crime. This is a

superset of arrests, because it includes people who are cited for a crime without actually

---

[7]These specifications include: all students on a four-day school week regardless of age, percentage of schools on a 4 day week, and various binary forms of these

93

being arrested, which is a common occurrence. The number of offenses is aggregated up to the county level, to match the level of treatment and the other dataset.

In addition to providing detailed data on individual crimes, an advantage of the NIBRS data is that it flags the exact date of the offense as well as detailed demographic characteristics of the offenders; this allows us to precisely identify juveniles in the sample and examine changes in crime by day of the week. Unfortunately, Colorado has only been fully participating in the NIBRS program since 1997, so the data is not available for the entire sample period. Additionally, while the data covers most of Colorado, there are some agencies (approximately 10% of agencies covering approximately 10% of the state's population) who do not participate in the program, reducing the number of counties in the NIBRS sample to 61.[8] Table 3.2 contains summary statistics about offenses from NIBRS. The outcomes (both arrests and offenses) are normalized by population and defined as crime (offenses or arrests) per 1000 population in the county. Note that property crimes are a general category consisting of larceny, breaking and entering, grand theft auto, and arson. Violent crimes consist of homicide, sexual assault, robbery, and assault. Drug crimes are a separate category.[9]

---

[8]While I am able to adjust sample by dropping agencies that do not fully participate in NIBRS, any remaining underreporting should cause the coefficient to be biased downwards, as the number of reported crimes would be lower than actual number of crimes in the county.

[9]Some incidents involve more than one offense (i.e. breaking and entering while in possession of illegal drugs). We count such incidents in both categories. While this results in some double counting of incidents, the alternatives are less palatable. Dropping all of these incidents results in the loss of much of the data. Some systems use a hierarchy such that an incident is categorized as its most "severe" offense type; however, these distinctions are often arbitrary and can result in severe under-counting of some kinds of crimes, especially drug offenses.

Additionally, through the remainder of the paper we limit our estimation sample to only rural counties leaving us with 47 treated categories. Rural in this context is defined as counties that are not part of a metropolitan statistical area (MSA).[10] Four day schools are primarily undertaken in rural areas and consequently other rural areas make the best control group. In the counties identified as urban, four-day schools make up a small fraction of the total student body (less that 5%) whereas they make up a quarter of students in rural areas. Additionally, no urban county has been fully treated, i.e. have all of their students on a four-day week, or even had more than 20% of their students treated.[11] Many of the schools on four-day school weeks in urban areas are actually in outlying rural areas of urban counties. To count these as "treated" seems suspect as these areas likely have a very different crime pattern. While we feel strongly that this is the correct way to estimate the effect of a four-day school week, we acknowledge all may not feel the same way. Consequently, Section C.1 contains versions of the models we use estimated with the entire sample. While there are some minor differences, the results are largely the same.

---

[10]The MSAs ommitted are Denver, Boulder, Greeley, Colorado Springs, Fort Collins, Pueblo, and Grand Junction. Other possible definitions of rural, such as population counts or only omitting the Denver MSA, have been considered and have little influence on the estimated results.

[11]This matters because we would like to identify the effect off of true changes in schedule (i.e. adoption) and not just noise from smaller changes in student population (which would change our treatment in small ways). The "big" changes due to adoption are concentrated among the rural counties. Note that we have estimated the models using the percentage of schools treated (which avoids noise in the treatment at the cost of ignore the size of the schools treated) and obtained similar results.

3.3.2 Results

We estimate the following fixed-effects model

$$y_{ct} = \beta_0 + \beta_1 T_{ct} + \beta_2 x_{ct} + \gamma_c + \delta_y + \varepsilon_{ct} \tag{3.1}$$

where variable T_{ct} is the percent of students in a county in a given year that are on a four-day school week schedule, x_{ct} are county year level covariates (unemployment, percent of students in county eligible for free lunch, race, and student/teacher ratio), γ_c is a county level fixed effect, δ_y is a year fixed effect, and ε_{ct} is the error term. The treatment variable T_{ct} is constructed by dividing the total number of students in a county who are on the four-day week in a given year by the total number of students in a county-year. When using the Colorado DPS arrests data we drop the year that a school first adopts this policy since the crime data is reported by calendar year (as opposed to academic year) and the initial year would only be partially treated.

Table 3.3 reports the point estimates associated with Equation 3.1. The results in columns 7 and 8 show a significant increase in juvenile larceny. Switching all students in a county to a four-day school week leads to 0.75 increase in juvenile arrests for larceny per 1000 population. This effect is large, approximately 66% of a standard deviation, although effects of this magnitude are not unusual in the literature.[12] It is not implausible that students who have an additional unsupervised day off during the week are more

---

[12]For example, Jacob and Lefgren (2003) find effects on property crimes that roughly range between 50% and 75% of a standard deviation.

likely to engage in shoplifting and other petty theft, while other property crimes such as arson and breaking and entering are not affected. This possibly accounts for the fact that property crimes show a positive coefficient that is not statistically significant.[13]. These results are replicated with the NIBRS data in Table 3.4. The estimated coefficients are largely the same, but do not attain statistical significance (at least partially attributable to the shorter sample). Across both datasets, the estimated effects of the percent of students on a four-day school week on other types of crime appear sensible in direction and magnitude.. A reduction in drug and violent crime is consistent with Jacob and Lefgren (2003) results, but again is imprecisely measured.

3.3.3 Non-linear Treatment

The models presented to this point assume that the effect of the treatment is linear in the percent of students treated. To address the possibility that the treatment is non-linear, Table 3.5 and Table 3.6 contain results from models where the treatment is transformed into discrete bins, treatment between 0% and 10%, treatment between 10% and 30%, treatment between 30% and 100%, and treatment being equal to 100%.[14] These variables are all compared to the omitted (untreated) group. In this case, both the arrests and

---

[13]Even though larcenies are the largest subset of property crimes, analysis (unreported) of the other component parts of larceny indicate no effect, which is likely responsible for this null result.

[14]These divisions are chosen because they roughly divide the treated counties into quartiles. We have also tested various parametric forms of introducing non-linearity such as imposing a quadratic treatment. These models differ little from the results with a linear treatment.

the NIBRS data indicate that the effects of the policies are still significant and that the magnitude of the effect increases with treatment intensity.

3.3.4 Day of the Week Results

Aggregating the NIBRS data to day of the week, county, year level allows us to see if the four-day school week policy has any effect on specific days. For example, incapacitation would imply that students commit more crimes on days that they are not in school and are therefore unsupervised. Alternatively, a concentration story would tell us that we should expect higher levels of crime on the four days students are in school, but lower crime on the day they have off.

We run the following specification:

$$y_{dct} = \beta_0 + \beta_1 T_{ct} + \beta_2 DayOfWeek_d + \beta_3 T_{ct} * DayOfWeek_d + \beta_4 x_{ct} + \gamma_c + \delta_y + \varepsilon_{ct}$$

(3.2)

where *DayOfWeek* is a set of dummy variables indicating day of the week (Sunday is the omitted group) which is then interacted with the treatment variable. A significant coefficient on the interaction term would indicate that the treatment variable differentially affects crime on each day of the week. However, the interaction term coefficients reported in Table 3.7 are small and statistically insignificant.[15] This may fit with the scenario

---

[15]This is one instance where including the urban counties seems to make a difference. Table C.5 does indicates that there are some cyclical patterns to crime over the week that do vary on treatment; however, it is unclear if this is caused by treatment or by including observations with vastly different patterns in crime.

where there are small or nonexistent effects of incapacitation on any given day, but the overall increase of free time leads students to get into trouble more frequently.

3.4 Additional Robustness Checks

The results above are robust to a variety of alternative specifications, including changing the treatment definition to be percent of schools instead of students and varying the definition of rural. Two additional robustness checks, a placebo test using adult crime and a check for the exogeneity of policy decisions, are examined in this section.

3.4.1 Adult Crime Placebo Test

To ensure that the treatment variable is not picking up some underlying change, such as changes in the local economy or law enforcement practices, we run identical models using adult (age 25+) crime as the outcome.[16] If juvenile crime is actually increasing in areas because of the treatment (the four-day week policy) then we should not expect to see a statistically significant relationship between the percent of juveniles on a four-day week in a given county-year and adult crimes per 1,000. Table 3.8 contains the results for adult arrests data and Table 3.9 contains the results for NIBRS data. Neither dataset indicates a significant effect for adult crime, supporting the idea that a four-day school

---

[16]Individuals ages 18-25 are left out of these models. Some 18 year olds are still in high school (and are therefore in the treated population) and those just out of high school are likely to socialize with treated individuals and may be indirectly affected by treatment.

week policy is uniquely impacting juveniles and is not a proxy for other unobserved

changes in the county.

3.4.2   Leads of the Treatment Variables

A concern with our empirical approach which would threaten the identification of

our estimates is if school districts are adopting this policy in response to the current

crime rate in the county. For instance, if schools in counties with high crime rates adopt

the four-day week policy as a means to reduce crime (and potentially are adopting other

crime reducing policies concurrently), then the internal validity of our results will be

jeopardized.[17] We check for this type of adoption behavior by including leads of the

treatment variable to Equation 3.1. We follow a similar approach to that of Gruber and

Hanratty (1995), Friedberg (1998), and Bedard and Do (2005) and run the following

model:

$$y_{ct} = \beta_0 + \beta_1 T_{ct} + \beta_2 x_{ct} + \lambda_1 \triangle T_{ct+1} + \lambda_2 \triangle T_{ct+2} + \gamma_c + \delta_y + \varepsilon_{ct} \qquad (3.3)$$

Note that the two lead terms $\triangle T_{ct+1}$ and $\triangle T_{ct+2}$ are the percent of students in a

four-day school in a county between year t=0 and t=1 and then between t=1 and t=2,

respectively. Thus, the estimated coefficients on these two lead terms represent the

relationship between the change in four-day school week adoption between years t/t+1

---

[17]Note that this assumes that the error term ε_{ct} is uncorrelated over time. If there are unobserved factors that affect crime at time *t* and are correlated with subsequent adoption, this approach may not be valid.

and t+1/t+2 and crime in that county. If high crime rates in a county lead to the adoption of a four-day school week, we should observe school districts responding to high crime rates by adopting the four-day week in the following year and thus the estimates on the lead terms λ_1 and λ_2 should be positive and significant. Table 3.10 contains the results for the NIBRS data and the estimates on the lead terms are statistically insignificant and, even though they are occasionally large, vary widely in sign indicating that schools likely are not adopting this policy as a response to crime rates.[18]

3.5   Conclusion

We demonstrate that the implementation four-day school weeks is associated with increases in property crime, especially larceny, while effects on other types of crime are less clear. These increase in larceny committed by juveniles seems to be due to an overall increase in crime. We find little evidence that the four-day school week policy affected the distribution of crimes across days of the week. These results are consistent with an incapacitation story, where juveniles who have more unsupervised time are more likely to commit crimes, and are present in two different datasets. Additionally, the effects persist across a variety of robustness checks. Overall, these results suggest that

---

[18]This test is not feasible in the arrests data because we drop the year of adoption due to it being partially treated.

policymakers considering school schedule changes should also be aware of the possible

impacts on criminal behavior.

Figure 3.1: Adoption of Four-Day School Week Policy Across Colorado Counties Over Time

Year: 1993

Year: 2009
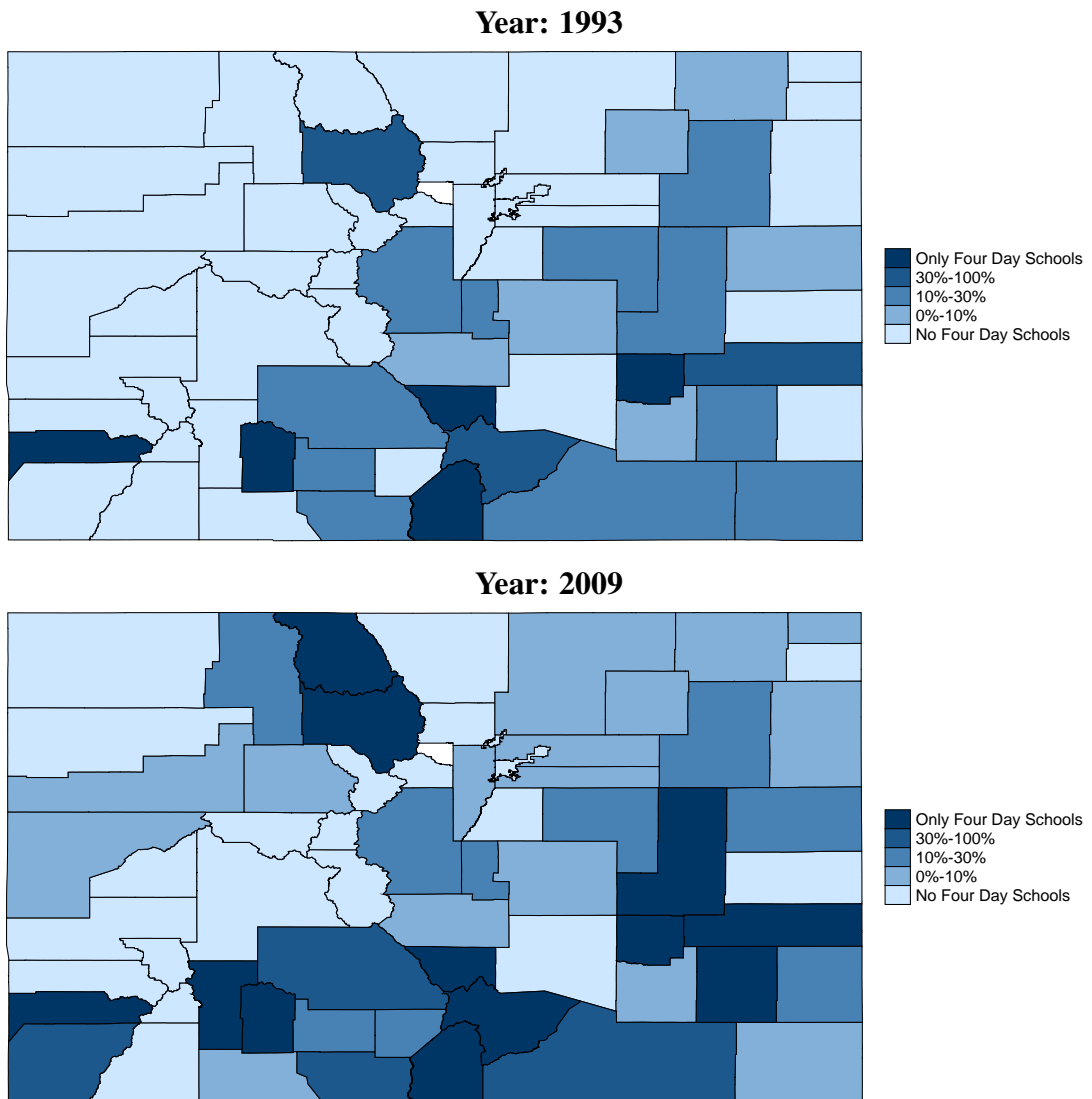
Table 3.1: Summary Statistics

	All Counties mean/sd	Urban Counties mean/sd	Rural Counties mean/sd	Untreated mean/sd	Treated mean/sd
Percent in 4-Day Schoolweek	0.187	0.0447	0.235		0.283
	(0.329)	(0.0894)	(0.365)		(0.370)
0 < Treatment < 0.1	0.148	0.140	0.151		0.225
	(0.356)	(0.348)	(0.358)		(0.418)
0.1 <= Treatment < 0.3	0.149	0.198	0.133		0.226
	(0.357)	(0.399)	(0.340)		(0.419)
0.3 <= Treatment < 1	0.0612	0	0.0818		0.0928
	(0.240)	(0)	(0.274)		(0.290)
Treatment=1	0.112	0	0.150		0.170
	(0.316)	(0)	(0.357)		(0.376)
Friday Off	0.551	0.572	0.544		0.835
	(0.498)	(0.496)	(0.498)		(0.372)
Monday Off	0.192	0.132	0.212		0.291
	(0.394)	(0.339)	(0.409)		(0.455)
Juv. Prop. Crime	1.507	2.503	1.171	1.471	1.518
	(1.653)	(2.021)	(1.354)	(1.515)	(1.719)
Juv. Vio. Crime	0.172	0.272	0.138	0.187	0.164
	(0.278)	(0.309)	(0.258)	(0.356)	(0.225)
Juv. Drug Crime	0.419	0.774	0.299	0.476	0.389
	(0.478)	(0.488)	(0.412)	(0.537)	(0.441)
Juv. Larceny	1.147	1.967	0.870	1.145	1.141
	(1.383)	(1.708)	(1.129)	(1.258)	(1.443)
Adult Prop. Crime	2.869	3.942	2.507	3.190	2.726
	(2.332)	(2.710)	(2.070)	(2.474)	(2.283)
Adult Vio. Crime	1.149	1.292	1.100	1.290	1.082
	(1.023)	(1.257)	(0.926)	(1.247)	(0.884)
Adult Drug Crime	2.108	2.982	1.814	2.563	1.908
	(2.036)	(2.349)	(1.829)	(2.269)	(1.923)
Adult Larceny	2.207	3.253	1.854	2.488	2.086
	(2.016)	(2.438)	(1.716)	(2.152)	(1.982)
Rural County	0.748	0	1	0.675	0.781
	(0.434)	(0)	(0)	(0.469)	(0.414)
Population	70351.1	239280.5	13416.5	80123.0	64836.5
	(139968.3)	(197936.4)	(12543.6)	(136226.8)	(141365.0)
Unemployment Rate	4.794	4.662	4.839	4.638	4.864
	(2.093)	(1.566)	(2.242)	(2.112)	(2.079)
Percent on Free Lunch	0.257	0.204	0.276	0.200	0.286
	(0.147)	(0.134)	(0.147)	(0.140)	(0.143)
Percent White	0.750	0.738	0.755	0.784	0.734
	(0.187)	(0.195)	(0.185)	(0.184)	(0.187)
Student/Teacher Ratio	0.0716	0.0574	0.0764	0.0694	0.0727
	(0.0173)	(0.00481)	(0.0173)	(0.0189)	(0.0165)
Observations	964	243	721	332	636

Rural counties are defined as those outside of a Metropolitan Statistical Area (MSA). Two counties contain both types, Monday off and Friday off, of four-day schools

Table 3.2: NIBRS Summary Statistics

	All Counties mean/sd	Urban Counties mean/sd	Rural Counties mean/sd	Untreated mean/sd	Treated mean/sd
Percent in 4-Day School Week	0.167	0.0407	0.211		0.244
	(0.312)	(0.0832)	(0.348)		(0.351)
0 < Treatment < 0.1	0.190	0.199	0.186		0.277
	(0.392)	(0.400)	(0.390)		(0.448)
0.1 <= Treatment < 0.3	0.136	0.182	0.120		0.198
	(0.343)	(0.387)	(0.325)		(0.399)
0.3 <= Treatment < 1	0.0552	0	0.0741		0.0806
	(0.228)	(0)	(0.262)		(0.272)
Treatment=1	0.100	0	0.135		0.147
	(0.301)	(0)	(0.342)		(0.354)
Friday Off	0.557	0.646	0.527		0.814
	(0.497)	(0.479)	(0.500)		(0.389)
Monday Off	0.208	0.138	0.232		0.304
	(0.406)	(0.346)	(0.422)		(0.460)
Juv. Prop. Crime	2.476	3.467	2.135	2.720	2.405
	(3.422)	(3.743)	(3.239)	(4.584)	(2.904)
Juv. Vio. Crime	0.842	1.239	0.705	0.902	0.827
	(1.203)	(1.313)	(1.132)	(1.565)	(1.048)
Juv. Drug Crime	0.744	1.201	0.587	0.809	0.719
	(1.206)	(1.475)	(1.055)	(1.316)	(1.154)
Juv. Larceny	1.204	1.747	1.017	1.365	1.163
	(2.005)	(1.958)	(1.989)	(2.916)	(1.566)
Adult. Prop. Crime	18.11	24.12	16.04	29.45	17.41
	(19.64)	(19.75)	(19.18)	(115.7)	(19.45)
Adult. Vio. Crime	3.514	4.067	3.324	4.864	3.189
	(3.186)	(3.650)	(2.991)	(8.583)	(2.653)
Adult. Drug Crime	1.481	1.881	1.343	2.814	1.368
	(1.500)	(1.486)	(1.482)	(11.79)	(1.365)
Adult. Larceny	9.141	11.22	8.424	19.56	8.232
	(11.49)	(9.418)	(12.04)	(98.00)	(9.916)
Rural County	0.744	0	1	0.700	0.758
	(0.437)	(0)	(0)	(0.459)	(0.429)
Population	74115.1	246118.0	14927.7	64273.3	78156.5
	(144690.5)	(203948.8)	(13500.3)	(117054.5)	(155460.8)
Unemployment Rate	4.672	4.579	4.704	4.515	4.731
	(1.849)	(1.579)	(1.933)	(1.887)	(1.828)
Percent on Free Lunch	0.260	0.187	0.285	0.188	0.292
	(0.147)	(0.120)	(0.147)	(0.136)	(0.141)
Percent White	0.742	0.758	0.736	0.801	0.716
	(0.177)	(0.169)	(0.179)	(0.162)	(0.177)
Student/Teacher Ratio	0.0715	0.0578	0.0762	0.0708	0.0716
	(0.0165)	(0.00457)	(0.0164)	(0.0205)	(0.0145)
Observations	707	181	526	227	484

Rural counties are defined as those outside of a Metropolitan Statistical Area (MSA). Two counties contain both types, Monday off and Friday off, of four-day schools

Table 3.3: Base Specification: Arrest Data from Colorado DPS

	Juv. Property		Juv. Violent		Juv. Drug		Juv. Larceny	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Percent in 4-Day Schoolweek	0.381 (0.520)	0.395 (0.542)	-0.062 (0.073)	-0.037 (0.076)	-0.187 (0.157)	-0.165 (0.159)	0.727*** (0.204)	0.753*** (0.206)
Unemployment Rate		-0.000 (0.043)		-0.020 (0.013)		-0.010 (0.011)		-0.025 (0.040)
Percent on Free Lunch		-1.678 (1.174)		0.989*** (0.275)		-1.520* (0.901)		-1.325 (0.790)
Percent White		-0.938 (2.147)		0.057 (0.328)		-1.082 (0.768)		-0.798 (1.471)
Student/Teacher Ratio		-2.709 (12.814)		-1.864 (2.052)		2.176 (1.954)		2.614 (5.193)
County Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	766	766	766	766	722	722	766	766
Mean Dep. Var.	1.205	1.205	0.140	0.140	0.299	0.299	0.888	0.888
S.D. Dep. Var.	1.399	1.399	0.258	0.258	0.411	0.411	1.142	1.142

The estimation sample is resticted to only rural counties (those outside an MSA). The outcome is measured as arrests per 1000 population and is observed annually. Since the treatment is administered on academic year, the initial year of treatment is dropped from the analysis.

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Standard errors clustered at the county level in parentheses.

Table 3.4: Base Specification: NIBRS

	Juv. Property		Juv. Violent		Juv. Drug		Juv. Larceny	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Percent in 4-Day School Week	0.601 (0.801)	0.777 (0.854)	0.008 (0.263)	-0.011 (0.283)	-0.503 (0.531)	-0.398 (0.513)	0.727 (0.436)	0.750 (0.454)
Unemployment Rate		-0.200 (0.200)		0.001 (0.043)		-0.016 (0.029)		-0.028 (0.073)
Percent on Free Lunch		2.873 (3.410)		1.854 (1.125)		-1.698 (1.106)		1.610 (1.670)
Percent White		6.921 (5.110)		1.237 (1.235)		-0.608 (0.964)		5.957** (2.622)
Student/Teacher Ratio		-18.171 (20.904)		-4.674 (4.702)		-5.815 (6.192)		-4.055 (8.680)
County Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	526	526	526	526	526	526	526	526
Mean Dep. Var.	2.135	2.135	0.705	0.705	0.587	0.587	1.017	1.017
S.D. Dep. Var.	3.239	3.239	1.132	1.132	1.055	1.055	1.989	1.989

The estimation sample is resticted to only rural counties (those outside an MSA). The outcome is measured as offenses per 1000 population and is observed by academic year.

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Standard errors clustered at the county level in parentheses.

Table 3.5: Nonlinear Treatment: Arrest Data from Colorado DPS

	(1) Juv. Prop. Crime	(2) Juv. Vio. Crime	(3) Juv. Drug Crime	(4) Juv. Larceny
0 < Treatment < 0.1	-0.022	0.066	-0.037	0.148
	(0.357)	(0.048)	(0.076)	(0.225)
0.1 <= Treatment < 0.3	-0.474	0.026	0.057	-0.225
	(0.399)	(0.040)	(0.069)	(0.277)
0.3 <= Treatment < 1	-0.311	-0.013	-0.090	0.067
	(0.447)	(0.080)	(0.112)	(0.250)
Treatment=1	0.198	0.009	-0.153	0.568**
	(0.489)	(0.082)	(0.156)	(0.259)
Time Vars.	Yes	Yes	Yes	Yes
County Fixed Effects	Yes	Yes	Yes	Yes
Year Fixed Effects	Yes	Yes	Yes	Yes
Observations	766	766	722	766
Mean Dep. Var.	1.205	0.140	0.299	0.888
S.D. Dep. Var.	1.399	0.258	0.411	1.142

The estimation sample is resticted to only rural counties (those outside an MSA). The outcome is measured as arrests per 1000 population and is observed annually. Since the treatment is administered on academic year, the initial year of treatment is dropped from the analysis.

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Standard errors clustered at the county level in parentheses.

Table 3.6: Nonlinear Treatment: NIBRS

	(1) Juv. Prop. Crime	(2) Juv. Vio. Crime	(3) Juv. Drug Crime	(4) Juv. Larceny
0 < Treatment < 0.1	0.516	0.325	-0.076	0.486
	(0.810)	(0.217)	(0.228)	(0.397)
0.1 <= Treatment < 0.3	0.983*	0.457**	0.083	0.567*
	(0.505)	(0.170)	(0.183)	(0.307)
0.3 <= Treatment < 1	1.340	0.165	-0.194	1.093**
	(1.032)	(0.421)	(0.218)	(0.511)
Treatment=1	0.920	0.102	-0.428	0.840
	(0.898)	(0.286)	(0.496)	(0.508)
Time Vars.	Yes	Yes	Yes	Yes
County Fixed Effects	Yes	Yes	Yes	Yes
Year Fixed Effects	Yes	Yes	Yes	Yes
Observations	526	526	526	526
Mean Dep. Var.	2.135	0.705	0.587	1.017
S.D. Dep. Var.	3.239	1.132	1.055	1.989

The estimation sample is resticted to only rural counties (those outside an MSA). The outcome is measured as offenses per 1000 population and is observed by academic year.

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Standard errors clustered at the county level in parentheses.

Table 3.7: Day of Week Analysis: NIBRS

	(1) Juv. Vio. Crime	(2) Juv. Drug Crime	(3) Juv. Prop. Crime	(4) Juv. Larceny
Percent in 4-Day	0.067	0.004	0.351*	0.276**
School Week	(0.055)	(0.089)	(0.175)	(0.119)
Monday	0.035*	0.032*	0.057	0.059
	(0.018)	(0.018)	(0.057)	(0.038)
Tuesday	0.077**	0.059**	0.005	0.061
	(0.031)	(0.025)	(0.057)	(0.045)
Wednesday	0.045*	0.077***	0.011	0.044
	(0.026)	(0.024)	(0.046)	(0.027)
Thursday	0.049**	0.062**	-0.032	0.034
	(0.023)	(0.023)	(0.052)	(0.029)
Friday	0.080**	0.111***	0.199**	0.101***
	(0.031)	(0.036)	(0.082)	(0.033)
Saturday	0.011	0.046**	0.060	0.061**
	(0.012)	(0.022)	(0.049)	(0.027)
(% in 4-Day)*Mon	-0.020	-0.061	-0.086	-0.049
	(0.030)	(0.051)	(0.075)	(0.052)
(% in 4-Day)*Tue	-0.050	-0.090	-0.005	-0.057
	(0.043)	(0.065)	(0.083)	(0.065)
(% in 4-Day)*Wed	-0.027	-0.047	-0.078	-0.040
	(0.056)	(0.035)	(0.074)	(0.041)
(% in 4-Day)*Thu	-0.030	-0.066	0.043	-0.038
	(0.037)	(0.063)	(0.074)	(0.039)
(% in 4-Day)*Fri	-0.084	-0.181**	-0.171	-0.042
	(0.054)	(0.072)	(0.133)	(0.068)
(% in 4-Day)*Sat	-0.009	-0.086	-0.064	-0.042
	(0.021)	(0.053)	(0.069)	(0.039)
Time Vars.	Yes	Yes	Yes	Yes
County Fixed Effects	Yes	Yes	Yes	Yes
Year Fixed Effects	Yes	Yes	Yes	Yes
Observations	2971	2971	2971	2971
Mean Dep. Var.	0.123	0.099	0.355	0.163
S.D. Dep. Var.	0.270	0.310	0.801	0.455

The estimation sample is resticted to only rural counties (those outside an MSA). The outcome is measured as offenses per 1000 population and is observed by academic year.

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Standard errors clustered at the county level in parentheses.

Table 3.8: Effect on Adult Crime: Arrest Data from Colorado DPS

	(1) Adult Prop. Crime	(2) Adult Vio. Crime	(3) Adult Drug Crime	(4) Adult Larceny
Percent in 4-Day Schoolweek	0.471 (0.383)	0.353 (0.332)	0.517 (1.074)	0.260 (0.205)
Unemployment Rate	-0.034 (0.060)	-0.054 (0.038)	-0.030 (0.060)	-0.001 (0.048)
Percent on Free Lunch	-1.405 (3.118)	0.564 (0.896)	-1.951 (1.538)	-0.893 (1.846)
Percent White	3.353 (3.219)	-1.183 (1.511)	-2.748 (3.089)	3.687 (2.538)
Student/Teacher Ratio	4.865 (14.141)	-11.850* (6.073)	-7.016 (10.698)	6.667 (11.727)
County Fixed Effects	Yes	Yes	Yes	Yes
Year Fixed Effects	Yes	Yes	Yes	Yes
Observations	766	766	721	766
Mean Dep. Var.	2.544	1.094	1.814	1.864
S.D. Dep. Var.	2.070	0.924	1.829	1.707

The estimation sample is resticted to only rural counties (those outside an MSA). The outcome is measured as arrests per 1000 population and is observed annually. Since the treatment is administered on academic year, the initial year of treatment is dropped from the analysis. Adults are defined as those age 25+.

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Standard errors clustered at the county level in parentheses.

Table 3.9: Effect on Adult Crime: NIBRS

	(1) Adult. Prop. Crime	(2) Adult. Vio. Crime	(3) Adult. Drug Crime	(4) Adult. Larceny
Percent in 4-Day School Week	0.960 (5.151)	-0.538 (0.663)	-0.136 (0.244)	3.132 (2.123)
Unemployment Rate	-0.148 (0.512)	-0.071 (0.101)	-0.104 (0.070)	-0.190 (0.270)
Percent on Free Lunch	8.643 (19.812)	0.001 (3.218)	-0.135 (1.568)	10.304 (12.528)
Percent White	70.480** (28.892)	-0.070 (4.196)	-0.230 (2.087)	53.035** (21.607)
Student/Teacher Ratio	-46.643 (61.527)	-5.279 (22.177)	8.936 (7.980)	-25.891 (46.467)
County Fixed Effects	Yes	Yes	Yes	Yes
Year Fixed Effects	Yes	Yes	Yes	Yes
Observations	526	526	526	526
Mean Dep. Var.	16.038	3.324	1.343	8.424
S.D. Dep. Var.	19.183	2.991	1.482	12.040

The estimation sample is resticted to only rural counties (those outside an MSA). The outcome is measured as offenses per 1000 population and is observed by academic year. Adults are defined as those age 25+.
* p < 0.1, ** p < 0.05, *** p < 0.01. Standard errors clustered at the county level in parentheses.

Table 3.10: Leads of Treatment: NIBRS

	Juv. Property		Juv. Violent		Juv. Drug		Juv. Larceny	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Percent in 4-Day School Week	1.352	1.839	0.135	0.143	-0.654	-0.988	1.024*	1.299*
	(1.100)	(1.272)	(0.317)	(0.362)	(0.682)	(0.840)	(0.590)	(0.694)
1-Year Lead in Adoption	0.597	0.691	-0.049	-0.080	-0.509	-0.825	0.545	0.575
	(1.220)	(1.410)	(0.408)	(0.498)	(0.499)	(0.679)	(0.545)	(0.653)
2-Year Lead in Adoption		0.191		0.045		-0.949		0.237
		(1.037)		(0.257)		(0.698)		(0.633)
Time Vars.	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
County Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	488	449	488	449	488	449	488	449
Mean Dep. Var.		2.302		0.769		0.642		1.124
S.D. Dep. Var.		3.303		1.196		1.122		2.121

The estimation sample is resticted to only rural counties (those outside an MSA). The outcome is measured as offenses per 1000 population and is observed by academic year.

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Standard errors clustered at the county level in parentheses.

Bibliography

Akee, Randall Q, Timothy J Halliday, and Sally Kwak, "Investigating the Effects of Furloughing Public School Teachers on Juvenile Crime in Hawaii," Technical Report 2013.

Anderson, D Mark, "In school and out of trouble? The minimum dropout age and juvenile crime," *Review of Economics and Statistics*, 2012, *96* (2), 318–331.

_ **and Mary Walker**, "Does shortening the school week impact student performance? Evidence from the four-day school week," *Andrew Young School of Policy Studies Research Paper Series*, 2012, (12-06).

_ , **Benjamin Hansen, and Mary Beth Walker**, "The minimum dropout age and student victimization," *Economics of Education Review*, 2013, *35*, 66–74.

Attanasio, Orazio, Britta Augsburg, Ralph De Haas, Emla Fitzsimons, and Heike Harmgart, "Group lending or individual lending? Evidence from a randomised field experiment in Mongolia," Technical Report 2011.

Banerjee, Abhijit, Arun G Chandrasekhar, Esther Duflo, and Matthew O Jackson, "The diffusion of microfinance," *Science*, 2013, *341* (6144).

_ , _ , _ , **and** _ , "Replication data for: The Diffusion of Microfinance V9 `http://hdl.handle.net/1902.1/21538`," 2014.

Banerjee, Abhijit Vinayak, "Microcredit Under the Microscope: What Have We Learned in the Past Two Decades, and What Do We Need to Know?," *Annu. Rev. Econ.*, 2013, *5* (1), 487–519.

Bedard, Kelly and Chau Do, "Are middle schools more effective? The impact of school structure on student outcomes," *Journal of Human Resources*, 2005, *40* (3), 660–682.

Berthelon, Matias E and Diana I Kruger, "Risky behavior among youth: Incapacitation effects of school on adolescent motherhood and crime in Chile," *Journal of Public Economics*, 2011, *95* (1), 41–53.

Beshears, John, James J Choi, David Laibson, Brigitte C Madrian, and Katherine L Milkman, "The effect of providing peer information on retirement savings decisions," Technical Report, National Bureau of Economic Research 2011.

Black, Sandra E, Paul J Devereux, and Kjell G Salvanes, "Staying in the Classroom and out of the maternity ward? The effect of compulsory schooling laws on teenage births," *The Economic Journal*, 2008, *118* (530), 1025–1054.

Bramoullé, Yann, Habiba Djebbari, and Bernard Fortin, "Identification of peer effects through social networks," *Journal of Econometrics*, 2009, *150* (1), 41 – 55.

Breza, Emily, "Peer effects and loan repayment: Evidence from the Krishna default crisis," Technical Report.

__ , **Arun G Chandrasekhar, and Horacio Larreguy**, "Mobilizing investment through social networks: evidence from a lab experiment in the field," Technical Report 2013.

Brock, William A and Steven N Durlauf, "Identification of binary choice models with social interactions," *Journal of Econometrics*, 2007, *140* (1), 52–75.

Brown, Jeffrey R, Zoran Ivković, Paul A Smith, and Scott Weisbenner, "Neighbors matter: Causal community effects and stock market participation," *The Journal of Finance*, 2008, *63* (3), 1509–1531.

Bursztyn, Leonardo, Florian Ederer, Bruno Ferman, and Noam Yuchtman, "Understanding mechanisms underlying peer effects: Evidence from a field experiment on financial decisions," *Econometrica*, 2014, *82* (3), 1197–1197.

Carrell, Scott E, Mark Hoekstra, and James E West, "Is poor fitness contagious?: Evidence from randomly assigned friends," *Journal of Public Economics*, 2011, *95* (7), 657–663.

Chandrasekhar, Arun and Matthew Jackson, "Tractable and consistent random graph models," Technical Report.

Chandrasekhar, Arun G. and Randall Lewis, "Econometrics of Sampled Networks," Technical Report 2012.

115

Chandrasekhar, Arun G, Cynthia Kinnan, and Horacio Larreguy, "Informal Insurance, Social Networks, and Savings Access: Evidence from a lab experiment in the field," Technical Report 2011.

_ , _ , **and** _ , "Information, networks and informal insurance: evidence from a lab experiment in the field," Technical Report 2011.

Conley, Timothy G, Christian B Hansen, Robert E McCulloch, and Peter E Rossi, "A semi-parametric Bayesian approach to the instrumental variable problem," *Journal of Econometrics*, 2008, *144* (1), 276–305.

Cox, G.W. and M.D. McCubbins, *Setting the Agenda: Responsible Party Government in the U.S. House of Representatives*, Cambridge University Press, 2005.

Cross, Amanda Brown, Denise C Gottfredson, Denise M Wilson, Melissa Rorie, and Nadine Connell, "The impact of after-school programs on the routine activities of middle-school students: Results from a randomized, controlled trial," *Criminology & Public Policy*, 2009, *8* (2), 391–412.

Donis-Keller, Christine and David L Silvernail, *Research brief: A review of the evidence on the four-day school week*, Center for Education Policy, Applied Research and Evaluation, University of Southern Maine, 2009.

Duflo, Esther and Emmanuel Saez, "The role of information and social interactions in retirement plan decisions: Evidence from a randomized experiment," *The Quarterly Journal of Economics*, 2003, *118* (3), 815–842.

Fowler, J. H., "Connecting the Congress: A Study of Cosponsorship Networks," *Political Analysis*, March 2006, *14* (4), 456–487.

Fowler, JH, "Legislative cosponsorship networks in the US House and Senate," *Social Networks*, 2006, *28*, 454–465.

Friedberg, Leora, "Did Unilateral Divorce Raise Divorce Rates? Evidence from Panel Data," *The American Economic Review*, 1998, *88* (3), 608–627.

Giné, Xavier and Dean S Karlan, "Group versus individual liability: Short and long term evidence from Philippine microcredit lending groups," *Journal of Development Economics*, 2014, *107*, 65–83.

_ , **Karuna Krishnaswamy, and Alejandro Ponce**, "Strategic default in joint liability groups: Evidence from a natural experiment in India," Technical Report 2011.

Giorgi, G. De, M. Pellizzari, and S. Redaelli, "Identification of social interactions through partially overlapping peer groups," *American Economic Journal: Applied Economics*, 2010, *2* (2), 241–275.

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

Graham, Bryan S, "Identifying social interactions through conditional variance restrictions," *Econometrica*, 2008, *76* (3), 643–660.

Grau, Elnabeth and Michael F Shaughnessy, "The Four Day School Week: An Investigation and Analysis," Technical Report 1987.

Griffith, Michael, "What Savings Are Produced by Moving to a Four-Day School Week?," Technical Report 2011.

Gruber, Jonathan and Maria Hanratty, "The Labor-Marker Effects of Introducing National Health Insurance: Evidence From Canada," *Journal of Business & Economic Statistics*, 1995, *13* (2), 163–173.

Herring, Chris, "School's New Math: The Four-Day Week," *Wall Street Journal*, 2010.

Hong, Harrison, Jeffrey D Kubik, and Jeremy C Stein, "Social interaction and stock-market participation," *The Journal of Finance*, 2004, *59* (1), 137–163.

_ , _ , **and** _ , "Thy neighbor's portfolio: Word-of-mouth effects in the holdings and trades of money managers," *The Journal of Finance*, 2005, *60* (6), 2801–2824.

Jackman, Simon, *pscl: Classes and Methods for R Developed in the Political Science Computational Laboratory, Stanford University* Department of Political Science, Stanford University 2012. R package version 1.04.4.

Jackson, Matthew O., *Social and Economic Networks*, Princeton University Press, 2008.

Jackson, Matthew O, Tomas Rodriguez-Barraquer, and Xu Tan, "Social capital and social quilts: Network patterns of favor exchange," *The American Economic Review*, 2012, *102* (5), 1857–1897.

Jacob, Brian A and Lars Lefgren, "Are Idle Hands the Devil's Workshop? Incapacitation, Concentration, and Juvenile Crime," *American Economic Review*, 2003, *93* (5), 1560–1577.

Kingsbury, Kathleen, "Four-Day School Weeks," *TIME Magazine*, 2008.

Krivitsky, Pavel N, Mark S Handcock, and Martina Morris, "Adjusting for network size and composition effects in exponential-family random graph models," *Statistical Methodology*, 2011, *8* (4), 319–339.

Lee, Lung-fei, "Identification and estimation of econometric models with group interactions, contextual factors and fixed effects," *Journal of Econometrics*, October 2007, *140* (2), 333–374.

__ **and Jihai Yu**, "Estimation of spatial autoregressive panel data models with fixed effects," *Journal of Econometrics*, 2010, *154* (2), 165 – 185.

__ **, Xiaodong Liu, and Xu Lin**, "Specification and estimation of social interaction models with network structures," *Econometrics Journal*, 2010, *13* (2), 145–176.

Lerner, Josh and Ulrike Malmendier, "With a little help from my (random) friends: Success and failure in post-business school entrepreneurship," Technical Report, National Bureau of Economic Research 2011.

Levitt, Steven D., "How Do Senators Vote? Disentangling the Role of Voter Preferences, Party Affiliation, and Senator Ideology," *The American Economic Review*, 1996, *86* (3), pp. 425–441.

Li, Shanjun, Yanyan Liu, and Klaus Deininger, "How important are endogenous peer effects in group lending? Estimating a static game of incomplete information," *Journal of Applied Econometrics*, 2013, *28* (5), 864–882.

Lin, Xu, "Identifying peer effects in student academic achievement by spatial autoregressive models with group unobservables," *Journal of Labor Economics*, 2010, *28* (4), 825–860.

Lochner, Lance and Enrico Moretti, "The Effect of Education on Crime: Evidence from Prison Inmates, Arrests, and Self-Reports," *The American Economic Review*, 2004, *94* (1), 155–189.

Luallen, Jeremy, "School's out... forever: A study of juvenile crime, at-risk youths and teacher strikes," *Journal of Urban Economics*, 2006, *59* (1), 75–103.

Manski, C F, "Identification of Endogenous Social Effects: The Reflection Problem," *The Review of Economic Studies*, 1993, *60* (3), 531–542.

Masket, Seth, "Where you sit is where you stand: The impact of seating proximity on legislative cue-taking," *Quarterly Journal of Political Science*, 2008, *3*, 301–311.

Moffitt, Robert A et al., "Policy interventions, low-level equilibria, and social interactions," *Social dynamics*, 2001, *4*, 45–82.

Parent, Olivier and James P. LeSage, "Spatial dynamic panel data models with random effects," *Regional Science and Urban Economics*, 2012, *42* (4), 727 – 738.

Peltzman, Sam, "Constituent Interest and Congressional Voting," *Journal of Law and Economics*, 1984, *27* (1), 181–210.

＿ , "An Economic Interpretation of the History of Congressional Voting in the Twentieth Century," *The American Economic Review*, 1985, *75* (4), pp. 656–675.

Poole, K.T. and H. Rosenthal, *Congress: A Political-Economic History of Roll Call Voting*, Oxford University Press, 2000.

Rogowski, Jon C and Betsy Sinclair, "Estimating the causal effects of social interaction with endogenous networks," *Political Analysis*, 2012, *20* (3), 316–328.

Sacerdote, Bruce, "Peer effects with random assignment: Results for Dartmouth roommates," *The Quarterly Journal of Economics*, 2001, *116* (2), 681–704.

Appendix A

Appendices to Chapter 1

A.1 Proofs

Proof. Identification in each individual network implies identification in the model with multiple networks. Assume peer effects are identified in the model for each individual network as in Equation 1.14. This implies that the matrices I_n, G_ℓ, and G_ℓ^2 are linearly independent. By the definition of linear independence, this means that the equation

$$c_1^\ell I_n + c_2^\ell G_\ell + c_3^\ell G_\ell^2 = 0$$

has only the trivial solution $c_1^\ell = c_2^\ell = c_3^\ell = 0$ for all ℓ in 1 to L. Now consider the block diagonal matrices I_{nB}, G^* and $(G^*)^2$ and note that ℓth element of the diagonal is the matrix outlined above I_n, G_ℓ, and G_ℓ^2. Since these matrices are linearly independent, the off-diagonal elements are all 0, and this is true for all ℓ, it follows that the block diagonal matrices are also linearly independent, and therefore peer effects are identified. □

Proof. Identification in the model with multiple networks does not imply identification in in any specific network. To satisfy this proof, it suffices to find a situation where peer effects are not identified in the subnetworks but are identified in the general model. I provide a simple case, where there are two subnetworks in the model, each where peer effects are not identified. If peer effects are identified in the general model, it means that the matrices I_{nB}, G^* and $(G^*)^2$ are linearly independent and the equation $d_1 I_{nB} + d_2 G^* + d_3 (G^*)^2 = 0$ has only the trivial solution. Suppose that the model for the subnetworks G_1 and G_2 is not identified, which means that the there exists at least one

nontrivial solution to the functions

$$c_1^1 I_n + c_2^1 G_1 + c_3^1 G_1^2 = 0$$

and

$$c_1^2 I_n + c_2^2 G_1 + c_3^2 G_2^2 = 0.$$

Suppose, by way of contradiction, that $c_1^1 = c_1^2$, $c_2^1 = c_2^2$, and $c_3^1 = c_3^2$ for at least one of the solutions to the models for G_1 and G_2. Because the block diagonal matrix consists of only these two outcomes and 0's on the off diagonal elements, if follows that these values yield a nontrivial solution to

$$c_1^1 I_{nB} + c_2^1 G^* + c_3^1 (G^*)^2 = 0$$

which means that the matrices are linearly dependent and peer effects are not identified in the general model, a contradiction. $\square$

Discussion: Note that if at least one of $c_1^1 \neq c_1^2$, $c_2^1 \neq c_2^2$, and $c_3^1 \neq c_3^2$ these values cannot provide a solution that makes the block diagonal system linearly dependent. Hence, peer effects can still be identified in the general model, even if they are not in the individual models. Note that a similar result will hold for a block diagonal matrix of any size which means that even if peer effects are identified in the general model, there may be any number of submodels where identification fails.

A.2 Cosponsorship Data

The cosponsorship network data was gathered and used by James Fowler (Fowler, 2006a,b) and is available for download at a website, `http://jhfowler.ucsd.edu/cosponsorship.htm`, maintained by the author with data for subsequent congressional sessions incorporated by Andrew Scott Waugh and Yunkyu Sohn. Cosponsorship relationships are derived not just from bills, but also resolutions, amendments, and procedural votes when applicable (all these are referred to as "bills" throughout the paper). Bills without cosponsors are dropped. Legislators who never sponsor or cosponsor a bill are also dropped from the analysis, to allow for row-weighting of the adjacency matrix. These individuals almost all serve for very short periods of time. The fact that more bills are introduced than ever reach a vote, combined with the inclusion of amendments and resolutions, means that the cosponsorship network is defined over a large number of bills for each session (sometimes over 10,000).

The roll call data is made available by Poole and Rosenthal (2000) at the website `voteview.org`. Roll call votes include a vote on any measure where a roll call vote is requested by members of the chamber and can include procedural measures (i.e. adjournment of the session) in addition to final votes on legislation. The data is processed to remove unanimous and near unanimous votes (votes with less than 3% dissenting) because they provide no information about the voting outcomes. Representatives who

do not vote in more than 50% of the roll-calls are dropped from the analysis due to small sample sizes. This is all done using the pscl package (Jackman, 2012), designed specifically for this purpose. Modifying the 3% and 50% thresholds has no impact on the results of the paper.

Appendix B

Appendix to Chapter 2

B.1   Questions Used to Elicit Networks

Individuals were asked about who they:

- borrow money from

- give advice to

- help with a decision

- borrow kerosene (or rice) from

- lend kerosene (or rice) to

- lend money to

- obtain medical advice from

- engage socially with

- are related to

- go to temple with

- invite to one's home

- visit in another's home

to form the friendship network. The primary network used in this paper is the union of responses to these questions. This limits the amount of measurement error in responses due to

Appendix C

Appendix to Chapter 3

C.1 Estimates with All Counties

This appendix contains tables with results from estimating the four-day school week models with the entire sample. The tables above are limited to only rural counties, as defined by the Census Bureau. As discussed in text, the results are largely similar, with a few exceptions. The first is that effects are more significant, likely due to the increased sample size. The second is that there are some patterns observed in the results for the estimation by days of the week and the adult placebo test that are somewhat different from the rural sample. The day of the week analysis shows that there may be some differences in the weekly cycle of crime due to treatment; the adult placebo test has a coefficient that nears (but does not attain) statistical significance at the 10% level on adult larceny. It is difficult to tell if these differences are actually due to effects of treatment or if they are due to substantially different patterns in crime across urban and rural counties.

Table C.1: Base Specification: Arrest Data from Colorado DPS (Whole Sample)

	Juv. Property		Juv. Violent		Juv. Drug		Juv. Larceny	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Percent in 4-Day Schoolweek	0.707	0.718	0.000	0.018	-0.177	-0.164	0.985***	0.983***
	(0.546)	(0.547)	(0.080)	(0.087)	(0.154)	(0.154)	(0.217)	(0.217)
Unemployment Rate		-0.036		-0.025*		-0.010		-0.057
		(0.050)		(0.013)		(0.010)		(0.044)
Percent on Free Lunch		-2.131*		0.774**		-1.505*		-1.553*
		(1.168)		(0.334)		(0.846)		(0.804)
Percent White		-1.253		-0.048		-1.238*		-0.776
		(2.635)		(0.511)		(0.720)		(1.924)
Student/Teacher Ratio		0.923		-1.132		2.031		4.434
		(13.125)		(2.132)		(1.880)		(5.753)
County Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	1025	1025	1025	1025	966	966	1025	1025
Mean Dep. Var.	1.560	1.560	0.176	0.176	0.419	0.419	1.181	1.181
S.D. Dep. Var.	1.713	1.713	0.285	0.285	0.479	0.479	1.417	1.417

The outcome is measured as arrests per 1000 population and is observed annually. Since the treatment is administered on academic year, the initial year of treatment is dropped from the analysis.

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Standard errors clustered at the county level in parentheses.

Table C.2: Base Specification: NIBRS (Whole Sample)

	Juv. Property		Juv. Violent		Juv. Drug		Juv. Larceny	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Percent in 4-Day School Week	0.512 (0.727)	0.607 (0.781)	0.118 (0.253)	0.103 (0.275)	-0.506 (0.526)	-0.418 (0.509)	0.592 (0.393)	0.605 (0.404)
Unemployment Rate		-0.171 (0.168)		-0.042 (0.039)		-0.038 (0.040)		-0.017 (0.063)
Percent on Free Lunch		2.757 (3.026)		1.387 (1.020)		-2.002* (1.037)		1.737 (1.587)
Percent White		7.696 (4.942)		0.598 (1.264)		-0.988 (0.945)		6.359** (2.481)
Student/Teacher Ratio		-11.459 (20.863)		-1.882 (5.277)		-3.160 (6.289)		-2.542 (8.305)
County Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	707	707	707	707	707	707	707	707
Mean Dep. Var.	2.476	2.476	0.842	0.842	0.744	0.744	1.204	1.204
S.D. Dep. Var.	3.422	3.422	1.203	1.203	1.206	1.206	2.005	2.005

The outcome is measured as offenses per 1000 population and is observed by academic year.

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Standard errors clustered at the county level in parentheses.

Table C.3: Nonlinear Treatment: Arrest Data from Colorado DPS (Whole Sample)

	(1) Juv. Prop. Crime	(2) Juv. Vio. Crime	(3) Juv. Drug Crime	(4) Juv. Larceny
0 < Treatment < 0.1	0.069	0.054	0.147*	0.205
	(0.266)	(0.044)	(0.077)	(0.190)
0.1 <= Treatment < 0.3	-0.376	0.036	0.126*	-0.155
	(0.368)	(0.041)	(0.068)	(0.252)
0.3 <= Treatment < 1	-0.068	0.003	-0.011	0.228
	(0.425)	(0.080)	(0.118)	(0.240)
Treatment=1	0.563	0.051	-0.075	0.817***
	(0.481)	(0.089)	(0.166)	(0.250)
Time Vars.	Yes	Yes	Yes	Yes
County Fixed Effects	Yes	Yes	Yes	Yes
Year Fixed Effects	Yes	Yes	Yes	Yes
Observations	1029	1029	970	1029
Mean Dep. Var.	1.555	0.176	0.419	1.177
S.D. Dep. Var.	1.711	0.285	0.478	1.416

The outcome is measured as arrests per 1000 population and is observed annually. Since the treatment is administered on academic year, the initial year of treatment is dropped from the analysis.

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Standard errors clustered at the county level in parentheses.

Table C.4: Nonlinear Treatment: NIBRS (Whole Sample)

	(1) Juv. Prop. Crime	(2) Juv. Vio. Crime	(3) Juv. Drug Crime	(4) Juv. Larceny
0 < Treatment < 0.1	0.093	0.238	0.121	0.325
	(0.545)	(0.190)	(0.160)	(0.297)
0.1 <= Treatment < 0.3	0.831*	0.475***	0.157	0.516*
	(0.459)	(0.175)	(0.176)	(0.276)
0.3 <= Treatment < 1	1.119	0.284	-0.131	1.021**
	(1.005)	(0.434)	(0.224)	(0.466)
Treatment=1	0.694	0.228	-0.414	0.710
	(0.807)	(0.298)	(0.508)	(0.457)
Time Vars.	Yes	Yes	Yes	Yes
County Fixed Effects	Yes	Yes	Yes	Yes
Year Fixed Effects	Yes	Yes	Yes	Yes
Observations	711	711	711	711
Mean Dep. Var.	2.506	0.851	0.748	1.228
S.D. Dep. Var.	3.528	1.236	1.208	2.094

The outcome is measured as offenses per 1000 population and is observed by academic year.

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Standard errors clustered at the county level in parentheses.

Table C.5: Day of Week Analysis: NIBRS (Whole Sample)

	(1) Juv. Vio. Crime	(2) Juv. Drug Crime	(3) Juv. Prop. Crime	(4) Juv. Larceny
Percent in 4-Day	0.073	0.005	0.304**	0.230***
School Week	(0.045)	(0.086)	(0.130)	(0.084)
Monday	0.045***	0.033***	0.079**	0.063***
	(0.013)	(0.012)	(0.036)	(0.023)
Tuesday	0.079***	0.071***	0.055	0.074***
	(0.019)	(0.023)	(0.036)	(0.027)
Wednesday	0.053***	0.069***	0.061*	0.071***
	(0.016)	(0.017)	(0.035)	(0.024)
Thursday	0.066***	0.067***	0.016	0.044**
	(0.018)	(0.018)	(0.032)	(0.017)
Friday	0.089***	0.108***	0.172***	0.089***
	(0.019)	(0.023)	(0.048)	(0.020)
Saturday	0.009	0.055***	0.084***	0.068***
	(0.007)	(0.016)	(0.030)	(0.017)
(% in 4-Day)*Mon	-0.036	-0.064	-0.121**	-0.057
	(0.026)	(0.045)	(0.051)	(0.035)
(% in 4-Day)*Tue	-0.054*	-0.108*	-0.074	-0.074*
	(0.029)	(0.061)	(0.060)	(0.042)
(% in 4-Day)*Wed	-0.038	-0.036	-0.146**	-0.079*
	(0.047)	(0.034)	(0.065)	(0.041)
(% in 4-Day)*Thu	-0.055*	-0.074	-0.022	-0.053**
	(0.033)	(0.060)	(0.051)	(0.025)
(% in 4-Day)*Fri	-0.098**	-0.178***	-0.139	-0.027
	(0.042)	(0.060)	(0.101)	(0.057)
(% in 4-Day)*Sat	-0.007	-0.094*	-0.094**	-0.052*
	(0.017)	(0.048)	(0.046)	(0.027)
Time Vars.	Yes	Yes	Yes	Yes
County Fixed Effects	Yes	Yes	Yes	Yes
Year Fixed Effects	Yes	Yes	Yes	Yes
Observations	4625	4625	4625	4625
Mean Dep. Var.	0.133	0.117	0.393	0.191
S.D. Dep. Var.	0.272	0.305	0.739	0.430

The outcome is measured as offenses per 1000 population and is observed by academic year.

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Standard errors clustered at the county level in parentheses.

Table C.6: Effect on Adult Crime: Arrest Data from Colorado DPS (Whole Sample)

	(1) Adult Prop. Crime	(2) Adult Vio. Crime	(3) Adult Drug Crime	(4) Adult Larceny
Percent in 4-Day	0.722*	0.501	0.721	0.581**
Schoolweek	(0.404)	(0.354)	(1.060)	(0.278)
Unemployment Rate	-0.043	-0.065*	-0.066	-0.021
	(0.056)	(0.035)	(0.063)	(0.046)
Percent on Free	-2.163	-0.034	-2.677	-1.698
Lunch	(2.814)	(1.129)	(1.611)	(1.727)
Percent White	1.850	-1.288	-4.005	2.286
	(3.604)	(1.989)	(3.189)	(3.015)
Student/Teacher	7.571	-8.413	-7.959	9.778
Ratio	(14.022)	(5.863)	(10.835)	(11.822)
County Fixed Effects	Yes	Yes	Yes	Yes
Year Fixed Effects	Yes	Yes	Yes	Yes
Observations	1025	1025	964	1024
Mean Dep. Var.	2.911	1.152	2.108	2.228
S.D. Dep. Var.	2.352	1.039	2.036	2.031

The outcome is measured as arrests per 1000 population and is observed annually. Since the treatment is administered on academic year, the initial year of treatment is dropped from the analysis. Adults are defined as those age 25+.

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Standard errors clustered at the county level in parentheses.

Table C.8: Leads of Treatment: NIBRS (Whole Sample)

	Juv. Property		Juv. Violent		Juv. Drug		Juv. Larceny	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Percent in 4-Day School Week	0.990	1.654	0.235	0.275	-0.617	-0.807	0.812	1.148*
	(0.985)	(1.194)	(0.319)	(0.377)	(0.676)	(0.859)	(0.519)	(0.627)
1-Year Lead in Adoption	0.312	0.469	0.022	0.017	-0.585	-0.837	0.376	0.421
	(1.152)	(1.328)	(0.407)	(0.501)	(0.498)	(0.672)	(0.519)	(0.619)
2-Year Lead in Adoption		0.189		0.138		-0.760		0.234
		(0.985)		(0.285)		(0.700)		(0.588)
Time Vars.	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
County Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	653	599	653	599	653	599	653	599
Mean Dep. Var.		2.653		0.915		0.800		1.306
S.D. Dep. Var.		3.488		1.275		1.255		2.124

The outcome is measured as arrests per 1000 population and is observed by academic year.

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Standard errors clustered at the county level in parentheses.