UNIVERSITY OF CALIFORNIA, SAN DIEGO
SAN DIEGO STATE UNIVERSITY

Social network effects on health: a multilayered approach

A dissertation submitted in partial satisfaction of the requirements
for the degree Doctor of Philosophy

in

Public Health (Global Health)

By

Holly B. Shakya

Committee in charge:

University of California, San Diego
Professor James H Fowler, Chair
Professor Anita Raj

San Diego State University
Professor Paula M Usita, Co-Chair
Professor Brian Finch
Professor Suzanne Lindsay

2012

The Dissertation of Holly B. Shakya is approved, and it is acceptable in quality and form for publication on microfilm and electronically:

Co-Chair

Chair

University of California, San Diego
San Diego State University
2012

DEDICATION

To my family, who are my essential social network, and my everyday reminder as to why it is I do this work

TABLE OF CONTENTS

ACKNOWLEDGEMENTS

I would like to thank Professor James Fowler for taking on a mathematically naïve individual such as myself, and offering me his insights into the vast and complex world of quantitative thinking. He willingness to share his nifty network reality with me opened the door to so many new and wonderful possibilities. Thank you James for your kindness and patience.

I would also like to acknowledge Professor Paula Usita for being a wonderful mentor and professional role model. My work in public health began with Paula and without her support and encouragement I would never have made it to this point.

I am in great debt to the rest of my dissertation committee: Professors Suzanne Lindsay, Anita Raj, and Brian Finch who have each offered me invaluable support and guidance along the way.

Chapter 1 has been accepted for publication in *Archives of Pediatrics and Adolescent Medicine* as Shakya HB, Christakis NA, Fowler JH, Parental influence on adolescent substance abuse. Chapter 2 is being prepared for publication as Shakya HB, Christakis NA, Fowler JH, Social network correlates of latrine ownership. Chapter 4 is being prepared for publication as Shakya HB, Christakis NA, Fowler JH. Network based detection of Indian village communities predicts latrine adoption better than socioeconomic or geographic measures.

Holly B Shakya is the primary author on all three papers.

VITA

Education

1991	BA (Philosophy) Pomona College
2012	Doctor of Philosophy in Public Health (Global Health) University of California San Diego / San Diego State University La Jolla/San Diego, California, USA

Professional Positions

2011-2012	Social networks analyst and consultant: Healthy Works Evaluation project (P.I. Dr. Suzanne Lindsay)
2011 – Present	Research assistant: UNICEF/UCSD joint project on the identification and measurement of social norms. University of California San Diego (P.I. Dr. Gerry Mackie)
2011 – Present	Research Assistant: Gates foundation social network research development team. University of California San Diego/Harvard University (P.I. Dr. Nicholas Christakis and Dr. James Fowler)
2009– 2011	Research Assistant: UCSD collaborative research fellowship recipient (P.I. Dr. James Fowler). University of California San Diego
2008-2009	Research and editorial assistant for Dr. Carleen Stoskopf, director of Graduate School of Public Health, San Diego State University
2007-2009	Research assistant for “Intergenerational Factors Affecting PSE Reduction in Grandparent Families” and “Child Health Outcomes in Grandparent Families”. (P.I. Dr. Paula Usita)

Teaching Experience

2009	Co-Instructor, PH101, Introduction to Public Health Graduate School of Public Health, San Diego State University

Publications

Shakya, HB, Christakis NA, Fowler, JH. Parental influence on substance abuse in adolescent social networks. *Archives of Pediatrics and Adolescent Medicine*. (forthcoming 2012)

Usita PM, Shakya HB, Aging families and immigration in Handbook of Families and Aging (forthcoming 2012)

Mackie, G, Shakya, HB. What are social norms? How are they measured?. Report for the UNICEF- UCSD Center on Global Justice Project Cooperation Agreement (2012)

Shakya, HB., Usita, PM., Eisenberg, C., Weston, J., & Liles, S. Family well-being concerns of grandparents in skipped generation families. *Journal of Gerontological Social Work;* 2012: 55: *39-54*

ABSTRACT OF THE DISSERTATION

Social network effects on health: a multilayered approach

by

Holly B. Shakya

Doctor of Philosophy in Public Health (Global Health)

University of California, San Diego, 2012
San Diego State University, 2012

Professor James H Fowler, Chair
Professor Paula M Usita Co-Chair

Background: Previous research suggests that both network contagion and connection contribute to social network effects. Interventions that integrate network-based approaches may be the most effective.

Objective: This dissertation aims to test network effects on two diverse outcomes in two diverse settings: substance abuse within high schools in the United States, and latrine ownership within villages in Southern India.

Methods: We analyzed data from two separate datasets. For Chapter 2 we used data from the National Longitudinal Study of Adolescent Health, a nationally representative multi-wave study that includes full social network data on adolescents in the United States. We examined the effect of friends' parents' authoritative parenting on substance abuse outcomes in adolescents. For Chapters 3 and 4 we used data from the Social Networks and Microfinance Study to test social network predictors of individual latrine ownership in rural South India. In Chapter 3 we examined the correlation between direct social contact latrine ownership and

individual latrine ownership. We also explored interactions of social contact latrine ownership with individual network position and demographic characteristics. In chapter 4 we compared the effects of social network latrine ownership proportions on individuals' latrine ownership at three levels: that of their direct social contacts, their network communities, and their villages.

Results: In Chapter 2, adolescents whose friends' parents were authoritative were significantly less likely to abuse substances, and this effect was only partially mediated through friends' substance abuse. In Chapter 3, the latrine ownership of individuals was significantly correlated with that of their direct social contacts. Furthermore, those more central to the network were more likely to own latrines although the direct effect of the latrine ownership of social contacts was also smallest for this group. In Chapter 4, while both the latrine ownership proportions of direct social contacts and of those within the village significantly predicted individual latrine ownership, the effect of the network community was the strongest.

Discussion: The results of this body of work provide support for incorporating network strategies into health behavior interventions, while highlighting some novel approaches beyond what has been traditionally used.

CHAPTER 1
Introduction

There is a growing body of literature on social ecology, which focuses on the reciprocal interactions between individuals and the various layers of their social environment[1] [2]. This shift in perspective reflects a deepening awareness of the complex dynamics that influence individual behavior. Social network analysis offers a powerful set of tools for operationalizing and testing the impact of these social dynamics. While social network analysis is not a theory per se, it operates on the assumption that relationships between human beings are crucial for understanding a wide variety of social processes. These social processes can substantially impact the contexts and behaviors that influence health outcomes, along many dimensions.

Brief introduction to network analysis

A social network is measured by eliciting the ties between people. As with any research topic, the population of interest can be defined in many ways: for instance we can look at a network of women of childbearing age in rural India, adolescents in high schools in the US, or the entire population of a neighborhood in New York City. Social network analysts use name generators as the primary survey tool for collecting data[3]. A name generator is a question asked of someone in order to elicit her social ties. An example of a commonly used name generator is “With whom do you discuss important matters?”. Depending upon the research question and the population of interest, we may use several name generators in a study or only one. We may analyze the networks created from each name generator separately, or we may combine them into one larger network.

Network effects

There are two main perspectives that can be used when studying the relationship between social networks and health. One is the idea of flow or contagion[4]. For instance in a seminal paper on the spread of health outcomes through a population, Christakis and Fowler (2007) were able to demonstrate that having a social contact who is obese increases a person's own chance of developing obesity[5]. This effect was not only significant for a person's own directly connected social contacts, but for connected contacts at up to three degrees of separation. That is, if A is connected to B who is connected to C who is connected to D, A's obesity significantly increases D's chance of becoming obese, even if A and D are not directly connected. Christakis and Fowler were able to demonstrate, by controlling for a number of possible confounding factors, that this effect was most likely the result of social contagion.

Determining social influence, however, is not a particularly easy task. It is difficult to differentiate between actual social influence, and *homophily*, which is the tendency for similar others to group together[6 7]. In other words, if a group of teens are friends, and they are all smokers, can we conclude that they are all smokers because of a peer influence effect or are they friends because they are all smokers? While measuring the smoking behavior of teens before and after friendship formation may answer that question to some degree, it becomes trickier when the similarities that drive the adoption of smoking behavior are related to qualities that make one prone to be a smoker[7]. For instance, adolescents with certain personality types may be more likely to smoke, and may also be more likely to form friendships together. The challenge is to determine whether an individual would have been likely to start smoking regardless of their friendship network, or whether the friendship network

itself actually influenced the individual to engage in a behavior that, under different circumstances, they would not have adopted.

A second important way that we can use networks to understand health behavior and outcomes is through a structural approach[8]. This approach looks at the network as a whole, and then attempts to understand individual behaviors and characteristics as a function of the individual's network position, and of qualities of the network itself. For instance, prior research has shown that more densely connected networks are strong bastions of social influence, constraining individual behavior and making it difficult to innovate or stray from group norms[9 10]. Kohler and colleagues (2001) showed that women in densely connected Kenyan village networks were more or less likely to use contraception consistent with the norm of their proximal social network[9]. In a study on child mortality in Bangladesh, the highly central women in village networks were less likely to utilize professional child-birth services then more peripheral women because those at the center of the network were the most constrained by the prevailing norm of giving birth at home[10].

While network centrality can constrain innovation, in other contexts it can facilitate it. Gayen's study in Bangladesh (2010) found that centrally positioned women and their proximal network members were more likely to use contraception than those at the network periphery[11]. The authors hypothesized that this was because centrally located women had the most exposure to contraception education and interventions, and the most exposure to other women who had adopted the use of contraception. Because centrality is often associated with adoption of innovations, network-based interventions have been tested which successfully identify highly central "peer leaders" who are trained to diffuse important interventions throughout their respective communities[12 13]. For instance Kincaid (2000) implemented a

network-based strategy to contraception education in which educational meetings were held at the homes of highly central women, specifically including their proximal network members. Results of the trial indicated that contraception use in the network increased fivefold over an intervention utilizing simple health educator home visits[12].

This study aims to expand upon current network research by examining several aspects of network analysis that have been neglected or understudied: multidimensional network relations, the interaction between network influence and network position, and the impact of network derived communities.

Overview of main analyses

Chapter two, "Parental influence on substance use in adolescent social networks", explores the question of influence within adolescent networks by testing whether friend's parents authoritative parenting significantly effects the substance abuse outcomes of adolescents. This paper utilizes the network data available from the National Longitudinal Study of Adolescent Health, in which adolescents named their 5 closest friends, and then responded to a large battery of questions, including their relationship with their parents and their substance abuse behaviors. Results showed a significant correlation between the positive "authoritative" parenting behavior of an adolescent's friends and a decreased risk of substance abuse for that adolescent. This paper is unique because it sets the stage for a more multidimensional approach to network influence studies. While we know that friends influence each other[14-20], particularly during adolescence, that influence does not occur in a social vacuum. External to the homogenous adolescent-adolescent dyad are social influences that impact adolescent behavior. One of the most important is the influence of the parents[21-24], and these parental influences seem to work not only

through the expected relationship of own-parent to own-child but from parent to someone else's child through the channel of the relationship between adolescents.

Chapter 3, "Social network correlates of latrine ownership" tests the effect of direct social network relationships and network structural characteristics on household latrine ownership. This analysis uses complete network data collected from 75 villages in rural India. Controlling for known predictors of latrine ownership in India, we test whether a person's latrine ownership can be successfully predicted by the 1) latrine ownership of that person's social contacts and 2) that person's network characteristics (such as being at the periphery or the center of the network). We also test for interactions such as whether the effect of a social contact's latrine ownership varies by caste similarity, and or by significant network predictors. The results of this paper show a significant correlation between the latrine ownership of an individual and the latrine ownership of their social contacts, providing evidence that social influence may possibly be a crucial component of latrine ownership decisions. And while those most central to the network are the most likely to own a latrine, they are also the least likely to have friends with latrines. This suggests that if a social effect is occurring within this population, it may be most prominent for those at the periphery. This is in contradiction to previous evidence supporting interventions primarily directed at those who are most central.

Chapter 4 "New perspectives on latrine ownership in India: the effect of social network communities" tests a newly developing methodology in network analysis: the detection of network derived communities. In this paper, we use a network based community detection algorithm to uncover mutually exclusive communities within 75 villages in rural India. We then create aggregate measures of possible predictors of latrine ownership, including group level latrine ownership proportions, at the village

and community levels. Using a random effects model to account for within-group correlations, we find that not only does community level latrine ownership proportion significantly predict individual latrine ownership, but the effect size is greater than that found at the level of direct social contacts, or at the level of the village. These results are interesting along two related dimensions. First, it suggests that interventionists attempting to promote latrine ownership cannot simply focus on geographic markers of community. In this study community was more successfully determined by analyzing the underlying network structure within the geographic unit of the village. Secondly, it shows, that network community detection methods can be a powerful tool to understand the social dynamics of health behavior in general. More research is needed to understand the potential impact of interventions which engage network-based communities.

Work on social network analysis highlights the importance of social interactions and the impact of those interactions along a wide variety of dimensions. It offers a methodological toolkit for understanding those interactions in multi-dimensional ways. This body of work contributes to the application of network science in global settings by advancing the understanding of network contagion and connection. Results from the three papers will advance the field for future network researchers, particularly those who work in developing world settings, where the interactions between directly connected individuals can be a particularly strong determinant of individual behavior.

References

1. Grzywacz JG, Fuqua J. The social ecology of health: leverage points and linkages. *Behavioral Medicine* 2000;26(3):101-15.

2. Stokols D. Social ecology and behavioral medicine: implications for training, practice, and policy. *Behavioral Medicine* 2000;26(3):129-38.

3. Marin A, Hampton KN. Simplifying the Personal Network Name Generator. *Field Methods* 2007;19(2):163.

4. Christakis NA, Fowler JH. Social Contagion Theory: Examining Dynamic Social Networks and Human Behavior. *Statistics in Medicine* 2012;in press.

5. Christakis NA, Fowler JH. The spread of obesity in a large social network over 32 years. *New England journal of medicine* 2007;357(4):370-79 %@ 0028-4793.

6. McPherson M, Smith-Lovin L, Cook JM. Birds of a feather: Homophily in social networks. *Annual Review of Sociology* 2001:415-44 %@ 0360-572.

7. Aral S, Muchnik L, Sundararajan A. Distinguishing influence-based contagion from homophily-driven diffusion in dynamic networks. *Proceedings of the National Academy of Sciences* 2009;106(51):21544-49.

8. Fujimoto K, Valente TW, Pentz MA. Network structural influences on the adoption of evidence-based prevention in communities. *Journal of Community Psychology* 2009;37(7):830-45 %@ 1520-6629.

9. Kohler HP, Behrman JR, Watkins SC. The density of social networks and fertility decisions: Evidence from South Nyanza District, Kenya. *Demography* 2001;38(1):43-58 %@ 0070-3370.

10. Gayen K, Raeside R. Social networks, normative influence and health delivery in rural Bangladesh. *Social Science & Medicine* 2007;65(5):900-14.

11. Gayen K, Raeside R. Social networks and contraception practice of women in rural Bangladesh. *Social Science & Medicine* 2010;71(9):1584-92.

12. Kincaid DL. Social networks, ideation, and contraceptive behavior in Bangladesh: a longitudinal analysis. *Social Science & Medicine* 2000;50(2):215-31.

13. Kelly JA. Popular opinion leaders and HIV prevention peer education: resolving discrepant findings, and implications for the development of effective community programmes. *AIDS care* 2004;16(2):139-50 %@ 0954-121.

14. Burt RD, Peterson AV. Smoking cessation among high school seniors. *Preventive Medicine* 1998;27(3):319-27.

15. Kaplan CP, Napoles-Springer A, Stewart SL, Perez-Stable EJ. Smoking acquisition among adolescents and young Latinas:: The role of socioenvironmental and personal factors. *Addictive Behaviors* 2001;26(4):531-50.

16. Chen PH, White HR, Pandina RJ. Predictors of smoking cessation from adolescence into young adulthood. *Addictive Behaviors* 2001;26(4):517-29.

17. Andrews JA, Tildesley E, Hops H, Li F. The influence of peers on young adult substance use. *Health Psychology* 2002;21(4):349.

18. Christakis NA, Fowler JH. The collective dynamics of smoking in a large social network. *New England Journal of Medicine* 2008;358(21):2249-58.

19. Rosenquist J, Murabito J, Fowler JH, Christakis NA. The spread of alcohol consumption behavior in a large social network. *Annals of Internal Medicine* 2010;152(7):426.

20. Mednick SC, Christakis NA, Fowler JH. The spread of sleep loss influences drug use in adolescent social networks. *PloS one* 2010;5(3):e9775.

21. Wood MD, Read JP, Mitchell RE, Brand NH. Do parents still matter? Parent and peer influences on alcohol involvement among recent high school graduates. *Psychology of Addictive Behaviors* 2004;18(1):19.

22. Baumrind D. Current patterns of parental authority. *Dev Psychol* 1971;4(1p2):1.

23. Maccoby EE, Martin JA. Socialization in the context of the family: Parent-child interaction. *Handbook of Child Psychology* 1983;4:1-101.

24. Steinberg L. We know some things: Parent-adolescent relationships in retrospect and prospect. *Journal of Research on Adolescence* 2001;11(1):1-19.

CHAPTER 2

Parental Influence on Substance Use in Adolescent Social Networks

Abstract

Objectives: Both peer and parental influences have been associated with the use of addictive substances in adolescence. We evaluated the relationship between the parenting style of an adolescent's peers' parents and an adolescent's substance use.

Design: Longitudinal survey

Setting: Adolescents across the United States were interviewed at school and at home

Participants: Nationally representative sample of adolescents in the United States

Main Exposure: Authoritative versus neglectful parenting style of adolescent's parents and adolescent's friends parents; adolescent substance use

Main Outcome Measures: Adolescent alcohol abuse, smoking, marijuana use, and binge drinking

Results: If an adolescent has a friend whose mother is authoritative, that adolescent is 40% (95% CI 12%-58%) less likely to drink to the point of drunkenness, 38% (95% CI 5%-59%) less likely to binge drink, 39% (95% CI 12%-58%) less likely to smoke cigarettes, and 43% (95% CI 1%-67%) less likely to use marijuana than an adolescent whose friend's mother is neglectful, controlling for the parenting style of the adolescent's own mother, school level fixed effects, and demographics. These results are only partially mediated by peer substance use.

Conclusion: Social network influences may extend beyond the homogeneous dimensions of own-peer or own-parent to include extra-dyadic influences of the wider

network. The value of parenting interventions should be re-assessed to take into account these spillover effects in the greater network.

Background and significance

Research on adolescent and adult social networks has focused on the impact of *peers* on risk behaviors involving drugs, tobacco, and alcohol use[1-8]. Networks may influence individual substance use behavior via the prevalence of substance use within the network as well as the interpersonal dynamics among network members[9,10]. These effects may have serious consequences; for example, the probability of a future overdose is related to both the number of members of an individual's social network using drugs and the degree of conflict within that network[11].

At the same time, there is evidence that *parents* may influence adolescents via their style of parenting[12-14]. The parenting styles framework encompasses four distinct parenting categories that are derived from two dimensions of interaction: (1) parental control (how much a parent intervenes in their adolescent child's life) and (2) parental warmth (how much positive affect a parent shows for their adolescent). *Authoritative* parents are warm and communicative, but they also exert appropriate control. *Neglectful* parents exhibit neither warmth nor control. *Authoritarian* parents exert control while lacking warmth, while *permissive* parents show warmth but do not exert control. Studies of these four parenting styles suggest that the authoritative parenting style is optimal, with long-term benefits including academic success, positive peer relationships, minimal delinquent behavior, risk avoidance, and positive psycho-social adjustment, including higher levels of psychological well-being[14-20]. Adolescents with authoritative parents are also less likely to have delinquent peer networks[21].

Here, we explore the possibility that parenting matters not only because of the direct and proximal effect of parent on child, but also because of the indirect and

more distal relationship between a parent and their adolescent child's friends. In other words, do the benefits of good parenting spill over, spreading from person to person and affecting multiple adolescents in a network? This question has implications both for how parents supervise the social networks of their adolescent children, as well as for how policy makers view the potential benefits of parenting education and interventions. In a previous cross-sectional study by Fletcher and colleagues, network authoritativeness (an average of the degree to which the parents of an adolescents' peers used authoritative parenting) was correlated with a decreased propensity towards delinquency, lower levels of substance abuse, and greater psychosocial competence[22]. To investigate this question more thoroughly using longitudinal analyses and complete network data, we use the National Longitudinal Study of Adolescent Health (*Add Health*), a source of data that contains information about adolescent social networks, their parents' styles of parenting, and self-reported measures of substance abuse. Using longitudinal dyadic network regression models, we measure the association between an adolescent's behavior and their friend's behavior, their mother's parenting style, and their friend's mother's parenting style.

Methods

Data

Add Health is a nationally representative study that explores multiple facets of adolescent well-being. Four waves of the Add Health study have been completed: Wave I was conducted in 1994-1995 and included adolescents who were then in grades 7th through 12th grade, Wave II in 1996, Wave III in 2001-2002, and Wave IV in 2007-2008. In Wave I of the Add Health study, researchers collected an "in-school"

sample of 90,118 adolescents chosen from a nationally-representative sample of 142 schools.

As part of the survey, these students named up to 5 male and 5 female friends who were later identified from school-wide rosters to generate information about each school's complete social network. A subset of this group was then chosen for in-depth follow-up in subsequent waves. This "in-home" sample was administered longer questionnaires about their social networks, health behaviors, family dynamics, and emotional/developmental outcomes. We drew our information about parenting and adolescent substance abuse from the Wave I and II in-home datasets.

Adolescent-friend dyads were included in each analysis only if the observations for both individuals included data on all measures of interest, and if the pair indicated that they were friends for both Wave I and Wave II. Furthermore, adolescents who indicated that they were siblings, either full or half were removed from the sample. Questions on maternal warmth were not asked of individuals for whom no one was acting in the role of mother (which could include non-biological mothers such as aunts or grandmothers). Table 2.1 provides summary statistics for the sample populations. Adolescents in our sample, compared to those in the complete AddHealth Wave II sample, were less likely to be black (13% vs. 23%), slightly less like to be Hispanic (13% vs. 17%), similar in likelihood to be Asian (8% vs. 7.4%), came from marginally less wealthy households (mean income 46,000 vs. 48.670) but had similar levels of parental education (mean value 5.62 vs. 5.45).

Measures

Adolescents in the Add Health dataset responded to a battery of questions regarding their parent's parenting behavior. Parental control was assessed using yes-no responses to seven questions from which we created a composite measure[20],

based on the average responses to all 7 questions (Cronbach's alpha of 0.63).[20] Adolescents whose parents were reported to exert below the median level of control are categorized as low control. Those above or equal to the median are categorized as high control. Maternal warmth was assessed using responses to five questions used in prior research[20]. Cronbach's alpha on the five questions was 0.85. Warmth, like control, was categorized by placing those at the median level of warmth and above in the high-warmth parenting category, and those below the median in the low-warmth parenting category. The combination of the control and warmth categories allows us to define four different parenting types[20] coded as follows: *Authoritative***:** high warmth, high control; *Authoritarian*: low warmth, high control; *Permissive*: high warmth, low control; *Neglectful*: low warmth, low control. While adolescent responses regarding their parents could be biased due to respondent error, Steinberg found that adolescent report was less biased than parent self-report as parents tend to err towards depicting their own behavior in the most positive light[23].

In a comprehensive section on substance use, adolescents were asked a variety of detailed questions about prior and current substance use, related to alcohol use, cigarette smoking, marijuana use, and binge drinking. We coded four separate dichotomous substance abuse outcomes from questions asked in Waves I and II to represent either having engaged in the behavior or not. For more details on variable coding please see the supplementary appendix (SA).

To identify the networks, we treated each friendship nomination as a "directed tie" from the namer to the named friend. We called interviewed individuals "adolescents" and the people that they named "friends". Dyadic observations were created so that each observation included data from both an adolescent and a friend at Waves I and II for adolescent-friend pairs observed in the data. Dyads in which the

adolescent and their friend were not friends in both Waves I and II were removed from the dataset. Likewise, we removed all adolescent-friend pairs for which data was missing for either the adolescent, the peer, or the peer's parent.

Controls variables included adolescent age, race (white, Hispanic, black, or Asian), and sex. We measured socioeconomic status with two separate variables: mother's self-reported education level, and mother's self reported household income. Because associations between peer's behaviors could be the result of neighborhood or other contextual factors relating to geographic proximity, we included school fixed effects in all models. This effectively eliminates any spurious correlations that may arise due to between-school variation in the incidence of the dependent variables.

While the total population for the AddHealth dataset was 20,746 for Wave I and 14,738 for Wave II, our final sample was much smaller due to our strict inclusion criteria and due to missing data on some measures. Also, our measure for SES included mother's education, a variable that was only available among a subset of observations for whom a parent survey was conducted, which served to significantly lower the total sample size. The total number of egos was 1386 while the number of dyads used in the analyses ranged from 2003 to 2066.

Human Subjects

The research was approved by the institutional review board at the University of California, San Diego.

Analyses

We conducted separate regression analyses for each substance abuse outcome. A logit form of a general estimating equation was used to analyze each model testing the behavioral outcome of the adolescent at wave 2 as a function of

friend's mother's parenting at wave 2, controlling for friend's mother's parenting at wave 1, adolescent's and friend's behavior at wave 1, adolescent's mother's parenting at both waves, gender, age, SES, and school level fixed effects (see SA). Both adolescent and friend parenting were coded as four-category variables, with neglectful parenting used as the reference category against which the other three categories are compared (for detailed methods please see SA).

We used generalized estimating equation (GEE) procedures to account for multiple observations of the same adolescent across ego-friend pairings and we assumed an independent working correlation structure for the clusters (See SA for results of alternate analysis clustering on alters). To explore possible causal pathways by which influence may occur, we also present the results of a mediation analysis in which we tested the hypothesis that friend's mother's parenting influences friend's behavior, which in turn has an effect on the adolescent's behavior. To do so, we followed the steps of testing for mediation laid out by Baron and Kenny[24], using the results of a *Sobel* test (for details please refer to SA) to determine significance. For significant mediators we calculate the proportion of the main effect that is mediated by dividing the indirect effect by the main effect.

The GEE regression models in the tables presented in the main text and SA provide parameter estimates in the form of beta coefficients, whereas the results reported in the text and in Figures 2.2-2.4 are in the form of risk ratios. The key coefficient in these models that measures the effect of influence is on the variable for friend's mother's Wave II parenting style. Risk ratios were calculated from predicted probabilities of substance abuse as a function of parenting style (changing it from 0 to 1) with 95 % confidence intervals estimated using 1.96 plus or minus the se and assuming all other variables are held at their means.

Results

In Figure 2.1, we show social network graphs that include parenting styles and substance abuse behaviors. These figures illustrate that behavior tends to cluster in the social network, and that adolescents who do not engage in substance abuse are often connected to authoritative parents via their friends, even if their own parents are not authoritative (as evidenced by the large green squares in the figure).

Statistically, we first studied the relationship between an adolescent's behavior and their friend's behavior, controlling for the parenting style of the adolescent's parent and the adolescent's friend's parent, plus fixed effects and demographics (Figure 2.2). The behavior of an adolescent's friend is significantly associated with the behavior of the adolescent, such that having a friend who drinks to the point of drunkenness increases the probability of the adolescent doing the same by 32% (95% C.I. 1%-72%), having a friend who is a smoker increases the probability of the adolescent smoking by 90% (95% C.I. 48%-141%), having a friend who smokes marijuana increases the probability of an adolescent smoking marijuana by 146% (95% C.I. 62%-271%), and having a friend who is a binge drinker increases the probability of adolescent binge drinking by 47% (95% C.I. 9%-96%). (These estimates are net of the baselines behavior of both parties.) SA tables 2.1-2.4 show the results of all the analyses for all 4 outcomes, where the beta coefficient on the row for friends Wave II substance abuse shows the relevant result.

We then looked at the direct effects of an adolescent's mother's parenting style on the adolescent's behavior, controlling for the adolescent's friend's mother's parenting style (Figure 2.3). If an adolescent has an authoritative parent, the probability of drinking to the point of drunkenness is reduced by 57% (95% C.I. 20%-77%) and the probability of smoking is reduced by 43% (95% C.I. 3%-66%). These

results are presented in Table 2.3 for variable “Own mother authoritarian Wave II” for all 4 outcomes.

Finally, we tested the hypothesized network effect of the mother of an adolescent’s friend (Figure 2.4). If an adolescent has a friend whose mother is authoritative, that adolescent is 40% (95% CI 12%-58%) less likely to drink to the point of drunkenness, 38% (95% CI 5%-59%) less likely to binge drink, 39% (95% CI 12%-58%) less likely to smoke cigarettes, and 43% (95% CI 1%-67%) less likely to use marijuana than an adolescent whose friend’s mother uses authoritative parenting, controlling for the parenting style of the adolescent’s own mother, school level fixed effects, and demographics. Furthermore, if an adolescent has a friend whose mother is authoritarian, that adolescent is 46% (95% CI 6%-54%) less likely to use marijuana than an adolescent who friend’s mother is neglectful. These results are presented in SA tables 2.1-2.4 and the variable of interest is: **Friend mother authoritative Wave II.** Surprisingly, the strength of association with the parenting style of an adolescent’s *friend’s* mother is of about the same magnitude as the association with the parenting style of the adolescent’s *own* mother for alcohol abuse and smoking (the Wald test of differences between coefficient for own mother and friend’s mother with significance at p<=.05 was insignificant in both cases), while the association is stronger for friend’s mothers than own mother for marijuana smoking and binge drinking.

We conducted a mediation analysis SA tables 2.1-2.4 to explore whether parents may have a direct effect on their children’s friends, or if this effect is indirect, resulting from the direct effect on their own children, which then spreads through the adolescent social network. The results suggest that 7.7% of the association between friend’s mother’s authoritative parenting and an adolescent’s alcohol abuse behavior may be explained by the influence that the friend’s mother may have on the friend’s

behavior which in turn may influence the adolescent's behavior. This proportion is 8.9% for marijuana use, and 7.0% for binge drinking. The results of the mediation analysis were insignificant for smoking behavior. In all cases, the association of the friend's mother's parenting style with the friend's behavior was significant, as was the association between the friend's behavior and the adolescent's behavior. Furthermore, as can be seen in the last three columns of each table, adding friend's behavior to the model significantly reduced the association between the friend's mother's parenting and the adolescent's behavior. Sobel tests were significant in all cases, with the exception of alcohol abuse (which at 1.80 is only slightly below the 1.96 level required for significance). Hence, in all cases, the majority of the effect of peer's parents is direct.

Discussion

Most research on social networks focuses on social influence in direct relationships. In other words, when considering adolescent behavior, we tend to focus on their peers and parents, assuming that influence spreads only from peer to peer or from family member to family member. We have discounted less obvious social influences, or pathways that bridge more heterogeneous dimensions of an adolescent's social network.

This study used longitudinal complete network data to show a positive correlation between the parenting practices of an adolescent's friends' parents, and the substance abuse outcomes of that adolescent. Our analyses demonstrate that if an adolescent has friends whose parents use "authoritative parenting", that adolescent is less likely to abuse alcohol, smoke, use marijuana, and binge drink. Our results are consistent with previous research that shows the influence of both peers and parents on adolescent substance abuse outcomes, although in this study we find

that the indirect influence of a peer's parents may be just as important, if not more so. Furthermore, our results show that while the pathway between a friend's parent and an adolescent is partially mediated through the behavior of the peer, this accounts for only a small proportion of the observed relationship.

A large body of literature has supported the idea that peers influence adolescent substance abuse mainly through the modeling of behavior, social norms around substance use, and overt offers to participate in the behavior[26,27]. However, results of a study by De Vries and colleagues[28,29] challenge the peer influence paradigm, suggesting that similarity in smoking behavior among adolescents is likely a function of friendship selection, and that parental smoking behavior is both a stronger predictor of smoking adoption than peer influence as well as a significant predictor of choosing smoking peers. Both peer influence and peer selection based upon shared attributes surely occur[30-34]. Here, we demonstrate that a peer's engagement in substance abuse is strongly correlated with an increased probability of the adolescent initiating that same behavior. By controlling for endogenous factors, that is the baseline behavior of both the adolescent and his/her peers, we reduce the likelihood that choosing substance-abusing peers is the driving force behind the peer effect we observe in the model.

The influence of a parent, on the other hand, has been studied from the dimension of behavioral modeling[29,35] (adolescents with substance abusing parents are more likely to abuse themselves), as well as from the perspective of parenting practices. These are two distinct (though possibly interacting) pathways of influence as the parenting practices of an adolescent's family appear to promote positive outcomes through the shaping of psychological resilience and emotional well being, rather than simply as the result of modeling specific behaviors[36]. These practices

empower the adolescent to make beneficial choices and engage in positive behavior along a wide variety of dimensions.

The results of our mediation analysis suggest that, to some degree, the influence of the positive parenting of a friend's mother on an adolescent may be mediated through the behavior of the friend. That is, positive parenting discourages substance abuse in adolescents, which then leads to reduced substance abuse in their friends. However, this is only part of the story. The mediation model did not account for the majority of the observed effect. This suggests that positive parenting may benefit an adolescent's friendship network either through a buffering effect via the adolescent's positive psychological outcomes and behaviors and/or a direct contact effect with the friends' parent. That is, adolescents may have frequent contact with their friends' parents and may therefore benefit directly from observing the positive parenting interactions that are taking place within those families. A second possibility is that having peers who are psychologically bolstered by good parenting benefits an adolescent through the interactions between them, independent of whether or not those peers are modeling substance abuse behaviors. A third possibility is that an adult who uses positive parenting behaviors with their own adolescent child is also able to act as an effective mentor for that child's friends. Research on mentoring has identified ways in which unrelated adults can positively influence adolescents along many dimensions [37] partially because as these unrelated adults are external to the normal adolescent-parent conflict[14], adolescents may feel freer to express needs and concerns they may not be able to express with their own parents[38]. Mentoring is most successful when the relationship is long-term, imbued with positive affect, and the mentor is able to offer some sort of instrumental

support[38,39]. Positive relationships with friends' parents may have multiple advantages consistent with this view of successful mentorship.

This study has limitations. The results may not be generalizable to all adolescents in the United States, as the final network cannot be weighted to be nationally representative. Moreover, self-report substance abuse measures may be subject to bias due to social desirability or inexact recall. However, unlike measures used in many social influence studies, the peer substance abuse measures in this study are not reported as conjecture by the adolescent, but directly reported by the friend regarding their own behavior.

Any association between adolescents' drug use and their friends' parents' parenting style is based on observational data, and as such it is possible that either (1) adolescents are influenced by the neglectfulness of their friends' parents, and this neglectfulness promotes drug use or (2) parents are influenced by their children's friends' drug use, which causes them to become more neglectful. Darling and colleagues note that adolescents seek out non-parental adult role models[38], suggesting that parents affect adolescents and not the other way around, but it is important to stress that the association we report here may be in part due to reciprocal influence.

Conclusion

There is a body of evidence to suggest that offering education on parenting can bolster parenting competence which in turn results in a wide variety of improved outcomes for adolescents[40-42]. The results of our research suggest that investments in such interventions may pay off not only through the direct connection between parent and child, but through the less obvious direction of parent to child to child's friends, as

well directly from parent to child's friend. As a consequence, we may be undervaluing the total benefit that parenting education has on adolescent populations.[43]

Table 2.1. Summary Statistics *N (Respondent)=1386 N (Friend)=1404*		
	Wave I Value	*Wave II value*
Drunk in last year, Respondent %	26	29
Drunk in last year, Friend %	29	31
Cigarette in last month, Respondent %	24	32
Cigarette in last month, Friend %	37	35
Marijuana use in last month, Respondent %	11	13
Marijuana use in last month, Friend%	14	16
Binge drinking in last year, Respondent%	26	30
Binge drinking in last year, Friend%	28	31
Neglectful parenting, Respondent%	24	28
Neglectful parenting, Friend%	25	33
Permissive parenting, Respondent%	22	30
Permissive parenting, Friend%	24	30
Authoritarian parenting, Respondent%	24	22
Authoritarian parenting, Friend%	23	18
Age (Respondent), mean (SD)	16.68 (1.48)	
Female%	51	
Household Income (1000s of Dollars), mean (SD)	48.67 (40.48)	
Parent's Education, mean (SD)	5.62 (2.31)	
Hispanic %	13	
Black %	13	
Asian %	8	

Note: Parent's education is a 10 item scale (0 = never went to school; 1 = 8th grade or less; 2 = more than 8th grade, but did not graduate from high school; 3 = went to a business, trade, or vocational school instead of high school; 4 = high school graduate; 5 = completed a GED; 6 = went to a business, trade or vocational school after high school; 7 = went to college, but did not graduate; 8 = graduated from a college or university; 9 = professional training beyond a 4-year college or university)

Table 2.2. Bivariate association between friend's mother's parenting style (Wave II) and adolescent risk behavior *

	Binge drinking in last year[a]	**Smoked in last month[b]**	**Was drunk in last year[c]**	**Used marijuana in last month[d]**
	RR (95% CI)	RR (95% CI)	RR (95% CI)	RR (95% CI)
Friend mother permissive	0.87 (0.74-1.02)	0.93 (0.79-1.08)	0.84* (0.71-0.99)	0.80 (0.61-1.05)
Friend mother authoritarian	0.65** (0.52-0.81)	0.82* (0.67-0.99)	0.7** (0.56-0.86)	0.92 (0.68-1.24)
Friend mother authoritative	0.49** (0.38-0.64)	0.64** (0.5-0.79)	0.46** (0.36-0.6)	0.46** (0.31-0.68)

*reference is neglectful.

[a]Consumed 5 or more drinks in a row at one time within last year n=2056
[b]Smoked cigarettes at least once in last month n=2033
[c]Been drunk or high on alcohol at least once in last year. n=2061
[d]Smoked or used marijuana at least once in last month n=2003
*p<0.05, **p<0.01

Table 2.3. Multivariate association between friend's mother's parenting style and adolescent risk behavior *

	Binge drinking in last year[a] N=2056	**Smoked in last month[b] N=2033**	**Was drunk in last year[c] N=2061**	**Used marijuana in last month[d] N=2003**
	RR (95% CI)	RR (95% C.I.)	RR (95% C.I.)	RR (95% C.I.)
Friend mother permissive Wave II	0.82 (0.6-1.11)	0.87 (0.66-1.14)	0.8 (0.58-1.09)	0.87 (0.59-1.3)
Friend mother authoritarian Wave II	0.66* (0.46-0.94)	0.84 (0.61-1.16)	0.82 (0.56-1.19)	1.12 (0.72-1.72)
Friend mother authoritative Wave II	**0.62* (0.41-0.95)**	**0.61* (0.42-0.88)**	**0.6* (0.41-0.89)**	**0.57* (0.33-0.99)**
Friend mother permissive Wave I	1.25 (0.87-1.81)	1.17 (0.84-1.61)	1.27 (0.89-1.81)	1.01 (0.63-1.63)
Friend mother authoritarian Wave I	0.92 (0.63-1.36)	1.14 (0.81-1.56)	0.93 (0.63-1.36)	0.95 (0.59-1.51)
Friend mother authoritative Wave I	1.03 (0.69-1.54)	1.43* (1.03-1.95)	0.92 (0.62-1.36)	1.05 (0.63-1.76)
Own mother permissive Wave II	0.7 (0.44-1.13)	0.49** (0.32-0.75)	0.72 (0.46-1.11)	0.55* (0.3-1)
Own mother authoritarian Wave II	0.72 (0.43-1.19)	1.06 (0.68-1.62)	0.56* (0.34-0.92)	0.87 (0.45-1.67)
Own mother authoritative Wave II	0.58 (0.31-1.08)	0.58* (0.34-0.97)	0.43* (0.23-0.8)	0.86 (0.4-1.83)
Own mother permissive Wave I	0.89 (0.53-1.48)	0.87 (0.54-1.37)	0.88 (0.55-1.42)	1.49 (0.75-2.97)
Own mother authoritarian Wave I	0.62 (0.37-1.02)	0.94 (0.6-1.44)	0.84 (0.52-1.37)	1.93 (0.96-3.89)
Own mother authoritative Wave I	0.6 (0.34-1.06)	1.31 (0.81-2.06)	0.58 (0.34-0.99)	1.32 (0.63-2.76)
Friend use Wave 1	1.7** (1.29-2.25)	1.53** (1.17-1.97)	1.81** (1.38-2.38)	2.92** (1.94-4.38)
Own use Wave 1	7.53** (5.17-10.88)	6.77** (5.33-8.27)	7.33** (5.17-10.26)	11.14** (6.5-19.08)
Deviance	287.04	307.82	285.26	181.49
Null Deviance	427.8	449.19	427.81	232.14

*reference is neglectful

[a]Consumed 5 or more drinks in a row at one time within last year

[b]Smoked cigarettes at least once in last month

[c]Been drunk or high on alcohol at least once in last year.

[d]Smoked or used marijuana at least once in last month

All models run controlling for respondent age, gender, race, mother's education, mother's income plus school level fixed effects.

*p<0.05, **p<0.01

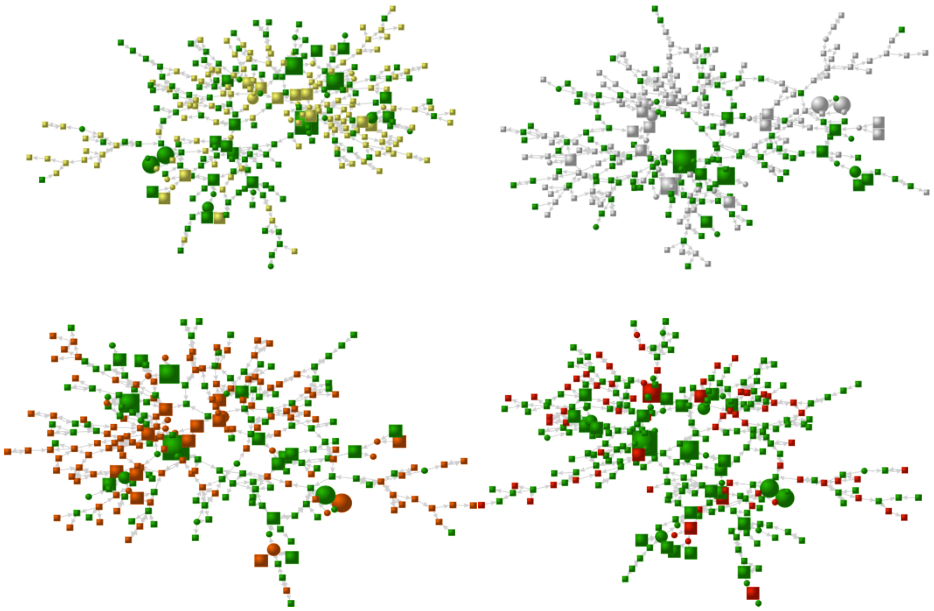

Figure 2.1. Illustrative network maps of one school in Add Health (*N*=304). Each node represents an adolescent and each arrow between them a friendship nomination. Node color indicates substance use behavior, yellow for drinking alcohol (upper left), gray for smoking tobacco (upper right), red for smoking marijuana (lower left), and orange for binge drinking (lower right). Green nodes indicate adolescents who do not engage in the substance abuse behavior shown in that panel. Circle nodes are adolescents with an authoritative parent, and square nodes are those with some other type (neglectful, authoritarian, or permissive). The size of each node is proportional to the number of friend's parents who are authoritative. These figures show that behavior tends to cluster in the social network, and adolescents who do not engage in substance abuse are often connected to authoritative parents via their friends, even if their own parents are not authoritative (indicated by large green squares).

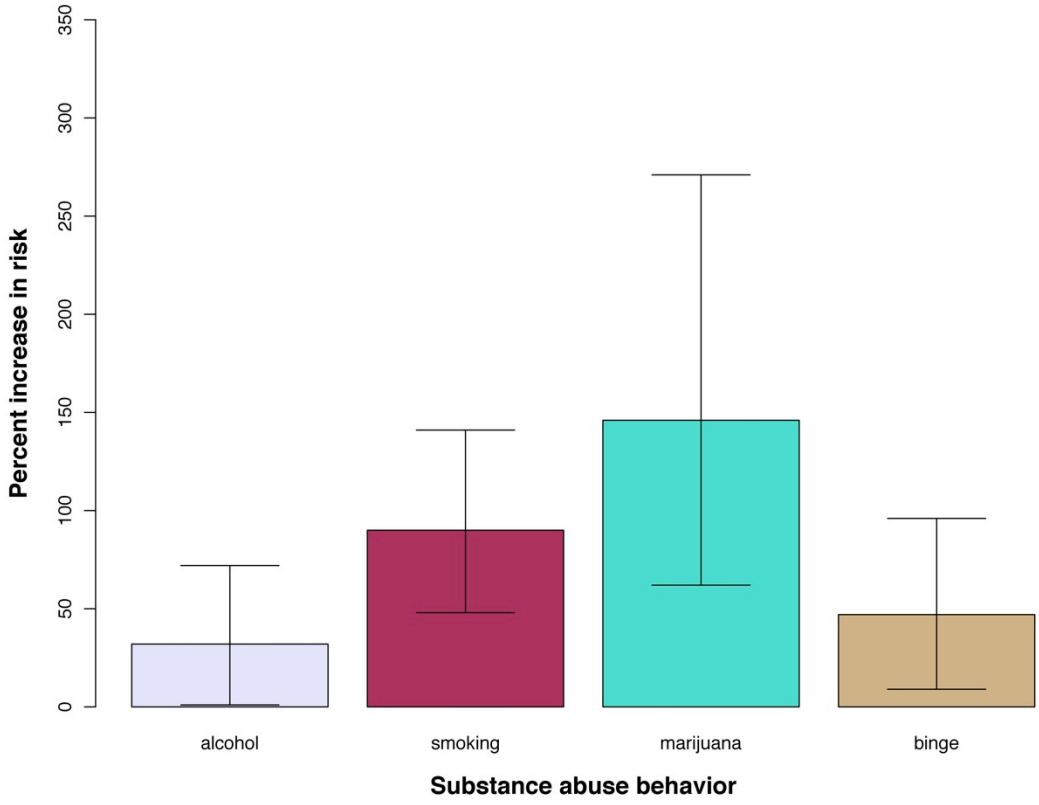

Figure 2.2. Percent increase in risk (includes 95% CI) of abusing alcohol, smoking, using marijuana, and binge drinking for an adolescent whose peer engages in the same behavior. All probabilities are estimated controlling for respondent age, gender, race, mother's education, mother's income, Wave I substance abuse, parent's Wave I and Wave II parenting style, friend's Wave I substance abuse, friend's parent's Wave I and Wave II parenting style, plus school level fixed effects.

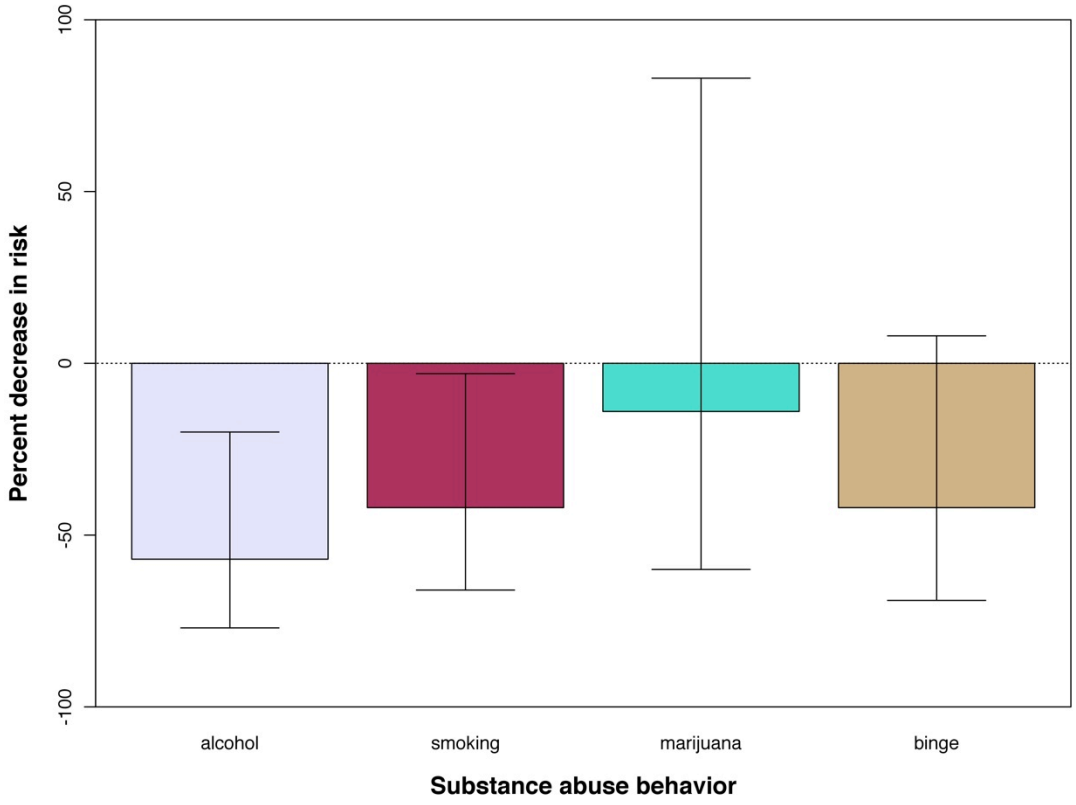

Figure 2.3. Percent decrease in risk (includes 95% CI) of abusing alcohol, smoking, using marijuana or binge drinking for adolescents whose parents are authoritative versus adolescents who parents are neglectful. All probabilities are estimated controlling for respondent age, gender, race, mother's education, mother's income, Wave I substance abuse, parent's Wave I parenting style, friend's Wave I substance abuse, friend's parent's Wave I and Wave II parenting style, plus school level fixed effects.

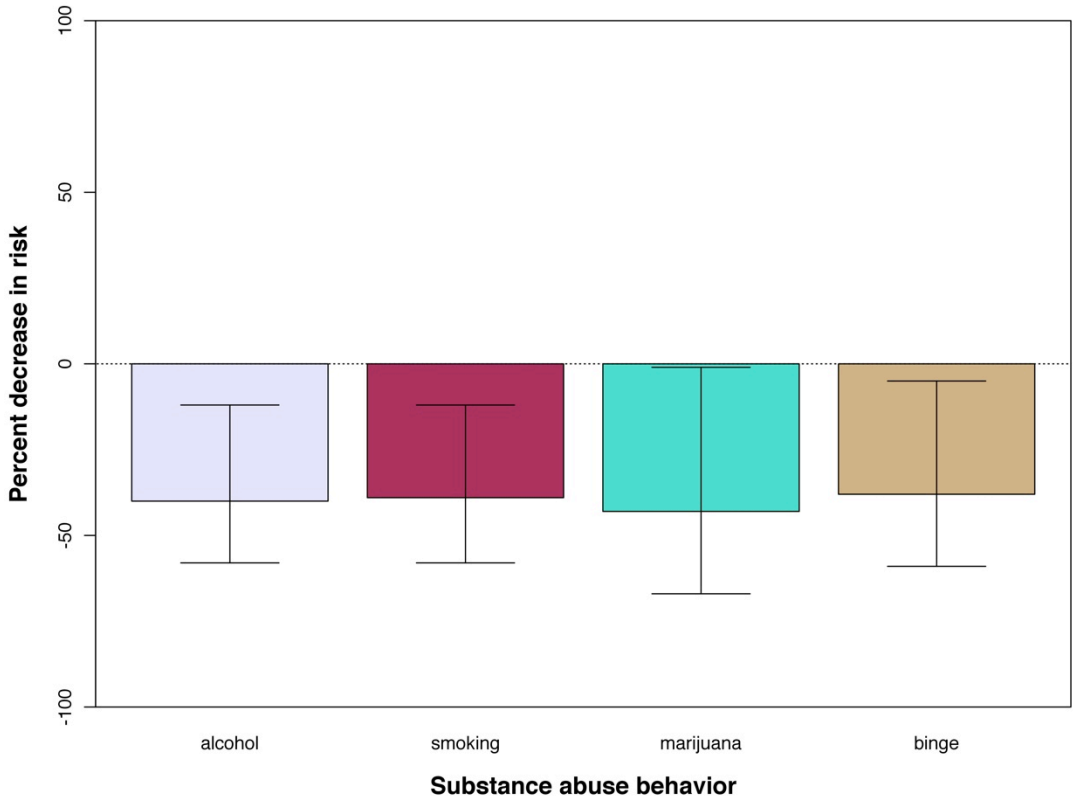

Figure 2.4. Percent decrease in risk (includes 95% CI) of abusing alcohol, smoking, using marijuana or binge drinking for adolescents whose peers' parents are authoritative versus adolescents whose peers' parents are neglectful. All probabilities are estimated controlling for respondent age, gender, race, mother's education, mother's income, Wave I substance abuse, parent's Wave I and Wave II parenting style, friend's Wave I substance abuse, friend's parent's Wave I parenting style, plus school level fixed effects.

References

1. Burt RD, Peterson AV. Smoking cessation among high school seniors. *Preventive Medicine.* 1998;27(3):319-327.

2. Kaplan CP, Napoles-Springer A, Stewart SL, Perez-Stable EJ. Smoking acquisition among adolescents and young Latinas:: The role of socioenvironmental and personal factors. *Addictive Behaviors.* 2001;26(4):531-550.

3. Chen PH, White HR, Pandina RJ. Predictors of smoking cessation from adolescence into young adulthood. *Addictive Behaviors.* 2001;26(4):517-529.

4. Andrews JA, Tildesley E, Hops H, Li F. The influence of peers on young adult substance use. *Health Psychology.* 2002;21(4):349.

5. Urberg KA, De*f*üirmencio*f*ülu SM, Pilgrim C. Close friend and group influence on adolescent cigarette smoking and alcohol use. *Dev Psychol.* 1997;33(5):834.

6. Christakis NA, Fowler JH. The collective dynamics of smoking in a large social network. *New England Journal of Medicine.* 2008;358(21):2249-2258.

7. Rosenquist JN, Murabito J, Fowler JH, Christakis NA. The spread of alcohol consumption behavior in a large social network. *Annals of Internal Medicine.* 2010;152(7):426.

8. Mednick SC, Christakis NA, Fowler JH. The spread of sleep loss influences drug use in adolescent social networks. *PloS one.* 2010;5(3):e9775.

9. Tyler KA. Social network characteristics and risky sexual and drug related behaviors among homeless young adults. *Social Science Research.* 2008;37(2):673-685.

10. Mundt MP. The Impact of Peer Social Networks on Adolescent Alcohol Use Initiation. *Academic Pediatrics.* 2011.

11. Latkin CA, Hua W, Tobin K. Social network correlates of self-reported non-fatal overdose. *Drug Alcohol Depend.* 2004;73(1):61-67.

12. Baumrind D. Current patterns of parental authority. *Dev Psychol.* 1971;4(1p2):1.

13. Maccoby EE, Martin JA. Socialization in the context of the family: Parent-child interaction. *Handbook of Child Psychology.* 1983;4:1-101.

14. Steinberg L. We know some things: Parent-adolescent relationships in retrospect and prospect. *Journal of Research on Adolescence.* 2001;11(1):1-19.

15. Demuth S, Brown SL. Family structure, family processes, and adolescent delinquency: The significance of parental absence versus parental gender. *Journal of Research in Crime and Delinquency.* 2004;41(1):58.

16. Steinberg L, Elmen JD, Mounts NS. Authoritative parenting, psychosocial maturity, and academic success among adolescents. *Child Dev.* 1989:1424-1436.

17. Gray MR, Steinberg L. Unpacking authoritative parenting: Reassessing a multidimensional construct. *J Marriage Fam.* Aug 1999;61(3):574-587.

18. Steinberg L, Mounts NS, Lamborn SD, Dornbusch SM. Authoritative parenting and adolescent adjustment across varied ecological niches. *Journal of Research on Adolescence.* 1991;1(1):19-36.

19. Bahr SJ, Hoffmann JP. Parenting style, religiosity, peers, and adolescent heavy drinking. *Journal of Studies on Alcohol and Drugs.* 2010;71(4):539.

20. Driscoll AK, Russell ST, Crockett LJ. Parenting styles and youth well-being across immigrant generations. *Journal of Family Issues.* 2008;29(2):185-209.

21. Knoester C, Haynie DL, Stephens CM. Parenting practices and adolescents' friendship networks. *J Marriage Fam.* 2006;68(5):1247-1260

22. Fletcher AC, Darling DE, Steinberg L, Dornbusch SM. The company they keep - relation of adolescents adjustment and behavior to their friends perceptions of authoritative parenting in the social network. *Dev Psychol.* Mar 1995;31(2):300-310.

23. Steinberg L, Lamborn SD, Darling N, Mounts NS, Dornbusch SM. Over time changes in adjustment and competence among adolescents from authoritative, authoritarian, indulgent, and neglectful families. *Child Dev.* 1994;65(3):754-770.

24. Baron RM, Kenny DA. The moderator‚Äìmediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations. *Journal of Personality and Social Psychology.* 1986;51(6):1173.

25. DuBois DL, Silverthorn N. Characteristics of natural mentoring relationships and adolescent adjustment: Evidence from a national study. *The Journal of Primary Prevention.* 2005;26(2):69-92.

26. Wood MD, Read JP, Mitchell RE, Brand NH. Do parents still matter? Parent and peer influences on alcohol involvement among recent high school graduates. *Psychology of Addictive Behaviors.* 2004;18(1):19.

27. Borsari B, Carey KB. Peer influences on college drinking: A review of the research. *Journal of Substance abuse.* 2001;13(4):391-424.

28. De Vries H, Candel M, Engels R, Mercken L. Challenges to the peer influence paradigm: results for 12-13 year olds from six European countries from the European Smoking Prevention Framework Approach study. *Tob Control.* 2006;15(2):83.

29. Biederman J, Faraone SV, Monuteaux MC, Feighner JA. Patterns of alcohol and drug use in adolescents can be predicted by parental substance use disorders. *Pediatrics.* 2000;106(4):792.

30. Christakis NA, Fowler JH. Social Contagion Theory: Examining Dynamic Social Networks and Human Behavior. http://arxiv.org/abs/1109.5235. 2011.

31. Hall JA, Valente T. Adolescent smoking networks: The effects of influence and selection on future smoking. *Addictive Behaviors.* 2007;32(12):3054-3059.

32. Hoffman BR, Sussman S, Unger JB, Valente T. Peer influences on adolescent cigarette smoking: A theoretical review of the literature. *Substance Use & Misuse.* 2006;41(1):103-155.

33. Valente T. Social network influences on adolescent substance use: An introduction. *Connections.* 2003;25(2):11-16.

34. Gaughan M. Predisposition and pressure: Mutual influence and adolescent drunkenness. *Connections.* 2003;25(2):17-31.

35. Pfistera RLASH, Wittchena HU. Maternal smoking and smoking in adolescents: a prospective community study of adolescents and their mothers. *Eur Addict Res.* 2003;9:120-130.

36. Collins WA, Maccoby EE, Steinberg L, Hetherington EM, Bornstein MH. Contemporary research on parenting - The case for nature and nurture. *Am Psychol.* Feb 2000;55(2):218-232.

37. Darling N, Hamilton SF, Shaver KH. Relationships outside the family: Unrelated adults. *Blackwell Handbook of Adolescence.* 2003:349-370.

38. Darling N, Hamilton SF, Niego S. Adolescents' relations with adults outside the family. 1994.

39. Forgatch MS, DeGarmo DS, Beldavs ZG. An efficacious theory-based intervention for stepfamilies*. *Behavior Therapy.* 2005;36(4):357-365.

40. Dishion TJ, Stormshak EA, Woody RH. Intervening in Children's Lives: An Ecological, Family-Centered Approach to Mental Health Care. *Adolescence* 2008;43(169):185.

41. Stormshak EA, Connell AM, V√©ronneau MH, et al. An ecological approach to promoting early adolescent mental health and social adaptation: family-centered intervention in public middle schools. *Child Dev.* 2011;82(1):209-225.

42. Christakis NA. Social networks and collateral health effects. *BMJ.* 2004;329(7459):184.

Supplementary appendix

Measures

Substance abuse measures

Adolescent history of alcohol abuse was assessed in Waves I and II by asking "Over the past 12 months, on how many days have you gotten drunk or "very, very high" on alcohol. A question on binge drinking asked "Over the past 12 months, on how many days did you drink five or more drinks in a row?" Tobacco use was measured by asking "During the past 30 days, on how many days did you smoke cigarettes?" Adolescent use of marijuana was assessed in Waves I and II by asking "During the past 30 days, how many times have you used marijuana?"

Parenting measures

Those without resident biological mothers responded to the questions regarding the woman performing the role of mother to them. Because questions regarding the father were less extensive than those regarding the mother, and at least 1/3 of the adolescents in the survey did not have a father in the home, our analysis is limited to responses regarding the adolescent's mother.

Parenting measures control

1.) "Do your parents let you make your own decisions about the time you must be home on weekend nights?" 2.) "Do your parents let you make your own decisions about the people you hang around with?" 3.) "Do your parents let you make your own decisions about what you wear?" 4.)"Do your parents let you make your own decisions about how much television you watch?" 5.)"Do your parents let you make your own decisions about which television programs you watch?" 6.)"Do your parents let you make your own decisions about what time you go to bed on week nights?" 7.)"Do your parents let you make your own decisions about what you eat?"

Parenting measures warmth

These were answered using a 5-point scale ranging from "strongly agree" (1) to "strongly disagree" (5): 1.)"Most of the time, your mother is warm and loving toward you." 2.) "Your mother encourages you to be independent." 3.)"When you do something wrong that is important, your mother talks about it with you and helps you understand why it is wrong." 4.)"You are satisfied with the way your mother and you communicate with each other." 5.) "Overall, you are satisfied with your relationship with your mother."

Model Specification

We specified longitudinal regression models with a basic form wherein the ego's status (e.g., smoking or not) at wave *t*+1, (Wave II) denoted y_{t+1}^{ego} (with distribution Y_{t+1}^{ego}), was a function of various time-invariant attributes of egos, such as gender and education (captured by the *k* variables denoted by *x* on the right), their status at wave *t* (Wave I), denoted (y_t^{ego}), and, most pertinently, the status of their alters at waves *t* (y_t^{alter}) and *t*+1 (y_{t+1}^{alter}).[1] This model is similar to those described by Valente[1].

We use generalized estimating equations to account for multiple observations of the same ego across waves and across ego-alter pairings[2]. And we only include observations in which ego and alter had a relationship at both wave *t* and wave *t*+1 on the grounds that people who are disconnected from each other should not influence each other that much, if at all[3].

Our basic model is thus:

$$g\left(E\left[Y_{t+1}^{ego}\right]\right) = \alpha + \beta_1 y_t^{ego} + \beta_2 y_{t+1}^{alter} + \beta_3 y_t^{alter} + \sum_{i=1}^{k} \gamma_i x_i \qquad (1)$$

1

where *g*() is a link function determined by the form of the dependent variable. For dichotomous data, $g(a)=\log(a / (1 - a))$.

Since we are using GEE, we also estimate the covariance structure of correlated observations for each ego. The covariance matrix of Y^{ego} is modeled by $V_{ego} = \phi A_{ego}^{1/2} R A_{ego}^{1/2}$ where ϕ is a scaling constant, *A* is a diagonal matrix of scaling functions, and *R* is the working correlation matrix. We assume an independence working correlation structure for the clustered errors, which has been shown to yield asymptotically unbiased and consistent, although possibly inefficient, parameter estimates (the β and γ terms) even when the correlation structure is mis-specified.[4]

To be clear, our basic model, based on an error correction model, assumes that there is no correlation of ego's substance use at *t*+1 with alter's substance use at *t*+1 except via influence, and no other effects on ego's substance use at *t*+1 except via the effect of ego's past substance use at wave *t* and the effect of the measured covariates (including ego's parents parenting style), i.e., conditional on no unobserved confounding[5]. These are common assumptions in regression models of observational data, of course. However, a special consideration here is that this assumption implies that there is no unobserved homophily beyond that on the observable variables.

The time-lagged dependent variable (lagged to the prior exam) helps control for ego's genetic endowment or any intrinsic, stable predilection to evince a particular trait. The lagged ego parenting variable helps control for homophily due to parenting styles, in other words the possibility that ego and alter were friends because of a similarity in their home environments. The lagged independent variable for an alter's trait (such as alter's parents parenting or alter's substance use) helps account for homophily (especially with respect to the observed trait that is the object of inquiry)

because it makes ego's current state unconditional on the state the alter was in when the ego and alter formed a connection[1] Conditioning on the lagged alter's trait, however, would not comprehensively deal with homophily on *unobserved* traits that are both time-varying and also associated with the outcome of interest. This term also does not address the issue of a shared context (confounding), but we deal with that by including a fixed effect for each adolescent's school. We have used this methodology to examine social influence in a variety of contexts[6-10]. For a review of this literature, see Fowler and Christakis (in press).[11]

Mediation model

A variable *M* mediates the relationship between an independent variable *X* and a dependent variable *Y* if (1) *X* significantly predicts *Y*, (2) X significantly predicts *M*, and (3) *M* significantly predicts *Y* controlling for *X* [12].While significant results for these tests do not prove there is a causal pathway (since the tests are based on observational data), they do allow us to reject the hypothesis of a causal relationship if the associations are not all significant. A formal test of mediation, called the Sobel test, determines whether the indirect effect is significantly different from zero[12 13]. The result of that Sobel test is a *z* score that can be compared to a critical value determined from a standard normal distribution. Like a *z* score, Sobel scores with a value greater than 1.96 are considered significant. Furthermore, mediation can be assessed by evaluating the degree to which the coefficient of the main effect of *X* on *Y* decreases in the presence of the mediator *M* (proposition 3 above). In complete mediation, the coefficient for *X* on *Y* should reduce to almost zero, meaning that the mediation path can explain the entire effect of *X* on *Y*. Partial mediation occurs when the coefficient of *X* on *Y* decreases in the presence of *M*, but not completely.

Supplementary Appendix References

1. Valente T. Models and methods for innovation diffusion. In: Carrington P, Scott J, Wasserman S, eds. *Models and Methods in Social Network Analysis*. Cambridge UK: Cambridge University Press; 2005.

2. Liang KY, Zeger SL. Longitudinal data analysis using generalized linear models. *Biometrika.* 1986;73(1):13-22 %@ 0006-3444.

3. Elwert F, Christakis NA. Wives and ex-wives: a new test for homogamy bias in the widowhood effect. *Demography.* 2008;45(4):851-873 %@ 0070-3370.

4. Schildcrout JS, Heagerty PJ. Regression analysis of longitudinal binary data with time-dependent environmental covariates: bias and efficiency. *Biostatistics.* 2005;6(4):633-652 %@ 1465-4644.

5. De Boef S, Keele L. Taking time seriously. *American Journal of Political Science.* 2008;52(1):184-200 %@ 1540-5907.

6. Christakis NA, Fowler JH. The spread of obesity in a large social network over 32 years. *New England journal of medicine.* 2007;357(4):370-379 %@ 0028-4793.

7. Christakis NA, Fowler JH. The collective dynamics of smoking in a large social network. *New England Journal of Medicine.* 2008;358(21):2249-2258.

8. Fowler JH, Christakis NA. The dynamic spread of happiness in a large social network. *BMJ: British medical journal.* 2008;337:a2338.

9. Rosenquist JN, Murabito J, Fowler JH, Christakis NA. The spread of alcohol consumption behavior in a large social network. *Annals of Internal Medicine.* 2010;152(7):426.

10. Mednick SC, Christakis NA, Fowler JH. The spread of sleep loss influences drug use in adolescent social networks. *PloS one.* 2010;5(3):e9775.

11. Christakis NA, Fowler JH. Social Contagion Theory: Examining Dynamic Social Networks and Human Behavior. *Statistics in Medicine.* 2012;in press.

12. Baron RM, Kenny DA. The moderator‚Äìmediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations. *Journal of Personality and Social Psychology*. 1986;51(6):1173.

13. Sobel ME. Asymptotic confidence intervals for indirect effects in structural equation models. *Sociological methodology.* 1982;13(1982):290-312.

Supplementary Appendix Tables

SAtable2.1 Mediation analysis testing whether the effect on adolescent alcohol abuse of the friend's parent's parenting style* is mediated through friend alcohol abuse

	Main effect: Friend's mother-> Own alcohol use		Main to mediator: Friend's mother->friend alcohol		Mediator to Outcome: Controlling for main predictor	
	Beta	SE	Beta	SE	Beta	SE
Intercept	-1.14	1.79	-2.53	1.45	-1.19	1.80
Friend mother permissive Wave II	-0.23	0.16	-0.32*	0.16	-0.22	0.16
Friend mother authoritarian Wave II	-0.20	0.20	-0.55**	0.18	-0.17	0.20
Friend mother authoritative Wave II	**-0.52***	**0.20**	**-0.88****	**0.22**	**-0.48***	**0.20**
Friend mother permissive Wave I	0.24	0.19	-0.28	0.18	0.25	0.19
Friend mother authoritarian Wave I	-0.08	0.20	-0.20	0.17	-0.07	0.20
Friend mother authoritative Wave I	-0.08	0.20	-0.33	0.19	-0.07	0.21
Own mother permissive Wave II	-0.34	0.23	0.14	0.17	-0.35	0.23
Own mother authoritarian Wave II	-0.60*	0.26	-0.35	0.20	-0.58*	0.26
Own mother authoritative Wave II	**-0.85**	**0.32**	**-0.33**	**0.21**	**-0.84***	**0.32**
Own mother permissive Wave I	-0.13	0.25	-0.04	0.18	-0.13	0.25
Own mother authoritarian Wave I	-0.18	0.25	-0.14	0.19	-0.17	0.25
Own mother authoritative Wave I	-0.56*	0.28	-0.02	0.19	-0.55*	0.28
Friend alcohol use Wave II					0.28*	0.14
Friend alcohol use Wave I	0.61	0.14	2.00**	0.13	0.50**	0.15
Own alcohol use Wave I	**2.08**	**0.19**	**0.34**	**0.14**	**2.07****	**0.19**
Deviance	285.26		305.61		283.80	
Null Deviance	427.82		449.72		427.82	
Sobel test for mediation -1.80						

N=2061

[a]Consumed 5 or more drinks in a row at one time within last year

* reference is neglectful

First three columns show logistic regression model of adolescent's past year alcohol abuse (the outcome variable) on friend's mother's parenting style(the explanatory variable). Second three columns show logistic regression model of friend's past year alcohol abuse (the mediator variable) on friend's mother's parenting style (the explanatory variable). Last three columns show logistic regression model of adolescent's past year alcohol abuse (the outcome variable) on friend's past year alcohol abuse (the mediator variable) controlling for friend's mother's parenting style (the explanatory variable). Models were estimated using a general estimating equation with clustering on the adolescent and an independent working covariance structure All models run controlling for controlling for respondent age, gender, race, mother's education, mother's income plus school level fixed effects.

SAtable2.2 Mediation analysis testing whether the effect on adolescent smoking of the friend's parent's parenting style* is mediated through friend smoking

	Main effect: Friend's mother-> Own smoking		Main to mediator: Friend's mother->friend smoking		Mediator to Outcome: Controlling for main predictor	
	Beta	SE	Beta	SE	Beta	SE
Intercept	-3.59	1.68	-2.95	1.28	-3.63	1.72
Friend mother permissive Wave II	-0.16	0.15	-0.33*	0.16	-0.14	0.16
Friend mother authoritarian Wave II	-0.19	0.18	0.27	0.19	-0.23	0.18
Friend mother authoritative Wave II	**-0.55***	**0.20**	-0.23	0.21	**-0.54***	**0.21**
Friend mother permissive Wave I	0.18	0.19	0.06	0.18	0.19	0.19
Friend mother authoritarian Wave I	0.15	0.19	-0.31	0.20	0.19	0.19
Friend mother authoritative Wave I	0.41	0.19	0.37	0.20	0.37	0.19
Own mother permissive Wave II	-0.79	0.23	-0.16	0.18	-0.79**	0.23
Own mother authoritarian Wave II	0.06	0.25	-0.01	0.20	0.06	0.26
Own mother authoritative Wave II	**-0.61**	**0.29**	-0.38	0.20	**-0.57***	**0.29**
Own mother permissive Wave I	-0.16	0.27	0.01	0.20	-0.17	0.27
Own mother authoritarian Wave I	-0.07	0.25	0.09	0.19	-0.09	0.25
Own mother authoritative Wave I	0.31	0.28	0.38	0.20	0.26	0.27
Friend smoking Wave II					0.73**	0.15
Friend smoking Wave I	**0.49**	**0.16**	**2.69****	**0.16**	0.10	0.17
Own smoking Wave I	**2.44**	**0.20**	**0.41***	**0.17**	**2.43****	**0.20**
Deviance	307.82		302.29		301.06	
Null Deviance	449.19		463.83		449.19	

Sobel test for mediation NS

N=2033

[a]Smoked cigarettes at least once in last month

*reference is neglectful

First three columns show logistic regression model of adolescent's past month smoking (the outcome variable) on friend's mother's parenting style (the explanatory variable). Second three columns show logistic regression model of friend's month smoking (the mediator variable) on friend's mother's parenting style (the explanatory variable). Last three columns show logistic regression model of adolescent's past month smoking (the outcome variable) on friend's past month smoking (the mediator variable) controlling for friend's mother's parenting style (the explanatory variable). Models were estimated using a general estimating equation with clustering on the adolescent and an independent working covariance structure. All models run controlling for controlling for respondent age, gender, race, mother's education, mother's income plus school level fixed effects.

SAtable 2.3 Mediation analysis testing whether the effect on adolescent marijuana use of the friend's parent's parenting style* is mediated through friend marijuana use[a]

	Main effect: Friend's mother-> adolescent marijuana use		Main to mediator: Friend's mother->friend marijuana use		Mediator to Outcome: Controlling for main predictor	
	Beta	SE	Beta	SE	Beta	SE
(Intercept)	-1.50	2.37	-3.35	1.63	-1.46	2.43
Friend mother permissive Wave II	-0.14	0.20	-0.72	0.21*	-0.06	0.21
Friend mother authoritarian Wave II	0.11	0.22	-0.02	0.23**	0.13	0.23
Friend mother authoritative Wave II	**-0.56***	**0.28**	**-0.71**	**0.27***	**-0.51**	**0.28**
Friend mother permissive Wave I	0.01	0.24	-0.11	0.23	0.02	0.25
Friend mother authoritarian Wave I	-0.05	0.24	0.04	0.23	-0.06	0.24
Friend mother authoritative Wave I	0.05	0.26	0.16	0.24	0.03	0.26
Own mother permissive Wave II	-0.60*	0.31	0.26	0.21	-0.66*	0.31
Own mother authoritarian Wave II	-0.14	0.34	0.56	0.23*	-0.21	0.34
Own mother authoritative Wave II	-0.15	0.39	-0.21	0.29	-0.15	0.39
Own mother permissive Wave I	0.40	0.35	0.04	0.25	0.41	0.35
Own mother authoritarian Wave I	0.66	0.36	-0.02	0.22	0.66	0.36
Own mother authoritative Wave I	0.28	0.38	0.27	0.26	0.25	0.38
Friend marijuana Wave II					**0.90****	**0.21**
Friend marijuana Wave I	1.07**	0.21	2.65**	0.19	0.66**	0.22
Own marijuana Wave I	2.41**	0.27	0.55*	0.23	2.40**	0.27
Deviance	181.48		203.68		177.57	
Null Deviance	232.14		264.09		232.14	
Alter's mother authoritative: Sobel test for mediation: -2.24						

N=2003

[a] Smoked or used marijuana at least once in last month

*reference is neglectful

First three columns show logistic regression model of adolescent's past month marijuana use (the outcome variable) on friend's mother's parenting style (the explanatory variable). Second three columns show logistic regression model of friend's month marijuana use (the mediator variable) on friend's mother's parenting style (the explanatory variable). Last three columns show logistic regression model of adolescent's past month marijuana use (the outcome variable) on friend's past month marijuana use (the mediator variable) controlling for friend's mother's parenting style (the explanatory variable). Models were estimated using a general estimating equation with clustering on the adolescent and an independent working covariance structure. All models run controlling for controlling for respondent age, gender, race, mother's education, mother's income plus school level fixed effects.

SAtable 2.4 Mediation analysis testing whether the effect on adolescent binge drinking of the friend's parent's parenting style[*] is mediated through friend binge drinking[a]

	Main effect: Friend's mother-> adolescent binge drinking		Main to mediator: Friend's mother->friend binge drinking		Mediator to Outcome: Controlling for main predictor	
	Beta	SE	Beta	SE	Beta	SE
(Intercept)	-0.68	1.68	-4.27	1.48	-0.60	1.67
Friend mother permissive Wave II	-0.20	0.16	-0.60**	0.16	-0.16	0.16
Friend mother authoritarian Wave II	**-0.42**	**0.18***	**-0.55****	**0.17**	**-0.39***	**0.19**
Friend mother authoritative Wave II	**-0.48**	**0.22***	**-0.91****	**0.22**	**-0.43***	**0.22**
Friend mother permissive Wave I	0.23	0.19	-0.41	0.18	0.26	0.19
Friend mother authoritarian Wave I	-0.08	0.20	-0.21	0.17	-0.07	0.20
Friend mother authoritative Wave I	0.03	0.21	-0.22	0.18	0.04	0.21
Own mother permissive Wave II	-0.35	0.24	-0.06	0.17	-0.35	0.24
Own mother authoritarian Wave II	-0.33	0.26	-0.40	0.19	-0.30	0.26
Own mother authoritative Wave II	-0.55	0.32	-0.30	0.20	-0.53	0.32
Own mother permissive Wave I	-0.12	0.26	0.13	0.19	-0.13	0.26
Own mother authoritarian Wave I	-0.49	0.26	0.13	0.19	-0.50	0.26
Own mother authoritative Wave I	-0.51	0.29	0.17	0.20	-0.52	0.29
Friend binge drinking Wave II					0.39*	0.15
Friend binge drinking Wave I	0.54	0.14**	1.87	0.14	0.40*	0.15
Own binge drinking Wave I	2.06	0.20**	0.29	0.14	2.05**	0.20
Deviance	287.04		312.56		284.87	
Null Deviance	427.80		440.02		427.80	
Sobel test for mediation: alter's mother authoritative -2.17						

N=2056

[a]Consumed 5 or more drinks in a row at one time within last year

* reference is neglectful

First three columns show logistic regression model of adolescent's past year binge drinking (the outcome variable) on friend's mother's parenting style (the explanatory variable). Second three columns show logistic regression model of friend's past year binge drinking (the mediator variable) on friend's mother's parenting style (the explanatory variable). Last three columns show logistic regression model of adolescent's past year binge drinking (the outcome variable) on friend's past year binge drinking (the mediator variable) controlling for friend's mother's parenting style (the explanatory variable). Models were estimated using a general estimating equation with clustering on the adolescent and an independent working covariance structure. All models run controlling for controlling for respondent age, gender, race, mother's education, mother's income plus school level fixed effects.

Chapter 3

Social network predictors of latrine ownership

Abstract

Background: Poor sanitation, including the lack of clean functioning toilets, is a major factor contributing to morbidity and mortality from infectious diseases in the developing world.

Methods: This study examines correlates of latrine ownership in rural India with a focus on social network predictors. Participants from 75 villages provided the names of their social contacts as well as their own relevant demographic and household characteristics. Using these measures, we tested whether the latrine ownership of an individual's social contacts is a significant predictor of individual latrine ownership. We also investigated whether network centrality significantly predicted latrine ownership, and if so, did it moderate the relationship between the latrine ownership of the individual and that of her social contacts.

Results: Our results showed that, controlling for the standard predictors of latrine ownership such as caste, education, and income, the latrine ownership status of an individual's social contacts significantly predicted their own latrine ownership. Interaction models show that this relationship is stronger among those of the same caste, the same education, and those with stronger social ties. We also found that higher measures of network centrality successfully predicted latrine ownership, and that centrality moderated the relationship between the latrine ownership of an individual and thather social contacts. The impact of the latrine ownership of a person's social contacts is higher for those who are less central in the network than for those who are more central.

Conclusion: Our results suggest that social effects may be a significant contributor to household latrine ownership decisions. Furthermore, interventions designed to promote latrine ownership should consider focusing on those at the periphery of the network, as they are less likely to own latrines while social spillover effects may be maximized in this part of the network.

Background and significance

The problem of sanitation

Diarrheal diseases are one of the most common causes of death for children under the age of 5, with almost 50% of those deaths occurring in India, Nigeria, Pakistan, Democratic Republic of Congo, and China [14]. Poor sanitation, including the lack of clean functioning toilets, is a major factor contributing to these outcomes [15] [16]. Exposed fecal matter pollutes ground water, drinking water, soil, and food sources [17]. Besides its contribution to morbidity and mortality due to diarrheal diseases, exposed fecal matter also spreads diseases like typhoid, cholera, and parasitic infections caused by worms [18]. Not surprisingly research shows that increasing access to latrines is a significant way to decrease diarrheal morbidity for children under 5 in India [19], although importantly the effect of a child's individual household latrine ownership is relatively insignificant compared to the effect of increasing community-wide latrine ownership. Increasing access to sanitation is an important part of the Millenium Development Goals (MDG), which were adopted by countries around the world as part of a united effort to improve the health and well being of the most world's most impoverished people [20].

Although India has made considerable progress in increasing the number of latrines around the country, close to 74% of the population still defecates in the open [17]. Rural access to latrines in the south Indian state of Karnataka, from which the data for this study was collected, is consistent with that of the rest of the country. While approximately 75% of urban households have access to latrines, only 17% of those in rural areas have access [21], although government sponsored latrine building campaigns in the last decade have increased that number to approximately 30% through some areas of the state [22].

Socio-economic predictors of household latrine ownership in India are not unexpected. Besides living in an urban area, those with the most education and highest incomes are the most likely to have latrines [17 23 24]. These factors affect both the resources to build a latrine as well as the exposure to, and understanding of, reasons why having a latrine would be beneficial. Caste may also serve as a proxy for many of these factors [17]. Those of the scheduled castes and scheduled tribe groups have historically been marginalized both socially and economically, and consistent with these characteristics they are the groups least likely to own latrines [17 24]. It has been suggested however, that these groups are also geographically stymied from latrine ownership due to housing situations with little access to the kind of drainage necessary for a properly functioning latrine [17]. In some areas it has also been found that Hindus are less likely to own latrines, as they believe that latrines situated close to their residences are polluting and find the latrines themselves "disgusting" [22].

More nuanced studies have found that even controlling for more obvious demographic factors, social norms and social expectations may drive latrine building decisions [16]. For instance, results from a qualitative study in Benin highlighted the fact that latrine building decisions were primarily rooted in motivations such as prestige, and had little to do with concerns regarding health [25]. Indian villagers have reported preferring open defecation because it allows people the chance to chat together, or because it is a time honored custom in their community [23]. Results from a study assessing the effects of Community Led Total Sanitation (CLTS) demonstrated that galvanizing entire communities was a necessary aspect of a latrine adoption campaign, and that the adoption decisions of the entire community were one of the

strongest influences in a household's decision to build a latrine [23]. The most successful latrine building campaigns have been those initiated by CLTS that successfully shift the norms of the community towards intolerance of open defecation, and community willingness to invest in building latrines [26]. While the Indian government has increased its efforts to provide latrine coverage for the nation, modeling its programs after CLTS, it has been unsuccessful at achieving positive results for those living in the most poverty, marginalized castes, and those in rural areas [17]. Government subsidies for building latrines have been largely unsuccessful as deeper social forces seem to affect the adoption of latrine usage within communities [27]. Pattanayak and colleagues (2009) have found, in fact, that for those above the poverty line, social shaming is a more economical and efficacious strategy for promoting latrine adoption than the use of subsidies, although monetary support may be necessary for those below the poverty line. [16]

Perspectives on social effects

Research on social norms has highlighted the difference between "descriptive norms", which are prevalent behaviors within a community, and "injunctive norms" or norms that are enforced within a community through sanctions: either positive sanctions for behaving within normative expectations, or negative sanctions for normative violations [30 31]. As new norms begin to become entrenched in a community, there is often a "tipping point", or a point at which a high enough proportion of the population has adopted the new process that from there on it begins to spread more rapidly, such that it may seem like an instantaneous change [32 33].

In some developing world communities the practice of open defecation may be simply a descriptive norm, and persist because it is commonly practiced. In other areas, however the practice may be an injunctive norm, and those who attempt to

transgress may be ridiculed or criticized. For instance, researchers in Kenya discovered that normative taboos around defecation include the belief that the feces of a father-in-law and daughter-in-law should not mix together, and therefore open defecation was a means by which this restriction was protected [34]. In order to discourage people from practicing open defecation, programs such as CLTS are attempting to foster injunctive norms to create an atmosphere of shame for those who defect from appropriate sanitation practices [26]. With this strategy, community members who defecate in the open receive negative sanctions, such as ridicule and public shaming. Villages are encouraged to be "open-defecation free", a process that requires all of the members of a community to change their behavior as well as monitor the behavior of others.

Social effects, a term which refers to the mechanism by which social behaviors can diffuse through a population, can differ according to the type of norm held in place. *Social learning* is more likely to take place in the context of descriptive norms, when individuals observe others engaging in a behavior which seems beneficial, such as having a latrine for their household [35 36]. The fact that others have taken the risk to engage in the behavior and seem to be accruing benefits from it makes it easier for the individual to adopt the behavior themselves. On the other hand, *social influence* processes are those that occur when behaviors are encouraged or constrained due to injunctive norms. The CLTS programs are designed with the idea that social influence dynamics can be cultivated in order to abolish the practice of open defecation within villages. Individuals are afraid to defect because to do so would risk the disapprobation of those in their social networks.

Social network analysis can be used to elucidate some of the nuances of these processes. Research on networks has suggested that many behavioral

processes are in fact the function of network dynamics. Broadly speaking, two main network mechanisms can impact the behavioral decision of any individual: connection and contagion.

Contagion occurs when information or behaviors spread through a network from individual to individual. Recent research on network contagion has suggested that a wide variety of health related behaviors and outcomes may spread from person to person through social networks [6 9 11]. These dynamics have been observed up to three degrees of separation, and they may also transcend homogeneous peer-to-peer network relations [37].

Network studies that look at connection offer additional insight by demonstrating how the structural position of an individual might impact that individual's behavior. Centrality measures, for instance, indicate which individuals are most connected within a network, and are positively correlated with their ability to influence others, and their tendency to be influenced [38-41]

Intervention strategies are increasingly focused on targeting the most central individuals with the idea that they will be able to positively influence others to adopt an innovation [42 43]. Banerjee and colleagues (2011), who collected the data we use in this article, studied the network diffusion of the decision to participate in a microfinance program among villagers in rural India. They found that the overall participation in microfinance programs was significantly higher when the introduction of information occurred among more centrally positioned individuals [44]. On the other hand, the most central individuals are also at a higher risk if the behavior or outcome being transmitted through the network is negative such as substance abuse, or an infectious disease.

Connection and contagion can interact in significant ways. For instance, Kohler and colleagues used network analysis to differentiate between social learning and social influence by showing that social influence tends to operate in denser networks where people are more closely connected and therefore more concerned with the behavior of others [35]. A study conducted in Bangladesh found that the most central women in the network were constrained from using professional birth services, presumably due to the strong normative influence of their overall network[45]. By contrast, social learning is more likely to occur in less densely connected networks in which individuals are less invested in the behavior of others, and where new ideas flowing between groups are more accessible[35].

This study

Few studies have quantitatively characterized the predictors of latrine ownership in developing country settings, and while the results of qualitative studies suggest that the success of latrine building campaigns seem to hinge on socially driven factors, even fewer studies have examined the relationship between social network characteristics and latrine ownership. A working paper by Dickinson and Pattanayak (2012) used the results of a randomized control trial of the CLTS program strategy to assess to what degree social factors drive latrine adoption. They found that accounting for a wide range of village and household characteristics, a household is more likely to adopt a latrine if their neighbors do. Similarly researchers in Benin found that neighborhood levels of latrine adoption were significantly associated with an increased probability that households in those neighborhoods would build a latrine [46].

While these studies are suggestive of network effects, none utilized social network data in which direct connections between individuals could be mapped and

analyzed. With social network data, connections between individuals are directly reported rather than inferred, allowing the calculation of direct rather than indirect effects between individuals and their social contacts: in other words we can statistically model the possibility of contagion. Of equal importance, having complete network data offers the opportunity to analyze the potential role of connection. Network centrality has been found to facilitate the spread of a new innovation in some contexts, while constraining that spread in others [45 47]. The analysis of network centrality offers researchers a powerful tool in the quest to understand the relationship between social dynamics and health behavior as well as the efficacy of health behavior interventions.

In this article we utilize network data collected from 75 villages in rural Karnataka to test whether having social contacts with latrines is associated with the probability that an individual will have a latrine. Going beyond these measures, we also test whether a person's place in the overall network is significantly associated with their probability of latrine ownership. While being central or peripheral to the network may impact an individual's chance of owning a latrine, potentially through increased exposure to latrine building norms or interventions, it may also act to moderate the relationship between the latrine ownership of an individual and that of their social contacts.

Methods

Data

We utilize *sociocentric* network data collected by Abhijit Banerjee and colleagues from 75 villages in rural Karnataka, in Southern India [44 48]. Sociocentric studies focus on a small population and attempt to ascertain all of the social relationships within a set of interconnected individuals[49]. This is in contrast to

egocentric network studies that focus on a larger population and attempt to ascertain all of the social relationships of a set of randomly-chosen individuals that are usually not connected to one another. Whereas egocentric data may help to improve the representativeness of a sample for a large population, sociocentric data allows measurement of larger network structures (like communities) and individual level network measures based on them.

In the Karnataka data, data was collected as part of a study to understand the network diffusion of micro-finance. A complete census was taken by interviewing one person within each household, including information regarding household characteristics such as latrine ownership, and roof construction. Individual surveys were used to collect demographic and network data from adult women from the ages of 18-57 and their eligible spouses from approximately half of the eligible households stratified by religion and geographic sub-location [44]. The total number of individuals interviewed was 16,984 of whom 16,608 were used in our analyses after those with missing observations were removed from the dataset.

Participants reported their age, gender, religion (Hindu or Muslim), and mother tongue (Kannada, Tamil, Telugu, or Hindi). Participants were also asked to identify to which caste they belong (scheduled caste, scheduled tribe, obc (other backward caste), or general). Education was measured using 16 levels ranging from none to higher degree. Household quality variables included roof type (6 categories), number of rooms in the home, number of beds in the home, and household electricity (private electricity, government electricity, or no electricity). Consistent with prior work in traditional agrarian societies in which data regarding actual income is unreliable, we used these household quality measures as a proxy for income [50]. An additional income measure was the type of ration card held by each individual. Ration cards are

used in India to guarantee government subsidies for food depending upon income, which we categorized as BPL (below poverty level), ABL (above poverty level), and not holding a card.

Our outcome variable was a binary measure of household latrine ownership.

Network measures

A name generator is the survey instrument used in social network data collection to elicit the important ties of individuals [51]. In this study respondents (termed here ***egos***) were asked to name up to 8 individuals (termed here ***alters***) for each name generator, and there were 12 name generators administered. These included asking respondents who they: 1) borrow money from, 2) give advice to, 3) help with a decision, 4) borrow kerosene or rice from, 5) lend kerosene or rice to, 6) lend money to, 7) obtain medical advice from, 8) engage socially with, 9) are related to, 10) go to temple with, 11) invite to one's home, 12) visit in another's home.

Because previous research has shown that networks derived from multiple name generators can more successfully measure network characteristics related to network composition [52], we collated data from all 12 name generators to create one comprehensive network. Each network tie indicated that one individual had named the other (or vice versa) in at least one name generator, and tie strength was calculated by summing the total number of name generators in which one of them named the other. The resulting data set yielded an undirected weighted network in which the weight on each tie ranged from 1 to 12.

Using the igraph library in R, we calculated three individual level network centrality measures for each individual in each village. *Degree centrality* [53] is simply the total number of unique social contacts that nominate or are nominated by the respondent. *Closeness centrality* [54] is the inverse of the average social distance

between a respondent and all other people in the network, where the social distance between friends is 1, between friends of friends is 2, and so on. To make its scale comparable to other variables, we transform it by multiplying it by 100. *Betweenness centrality* [55], is the number of shortest paths in the network that pass through a respondent divided by the total number of shortest paths. This measure identifies the extent to which an individual in the network is critical for passing support from one individual to another. The higher this number the greater the effect would be on the total average distance for the network if this person were removed [56]. Because of the skewed distribution of betweenness, we transform it by adding 1 and taking the log.

Homophily is the tendency for individuals to form social ties with others that are similar to them [57]. In this case we wanted to test whether the correlation between ego and alter's latrine ownership varied according to homophily on significant characteristics. We created binary measures for each dyad: one to assess whether the respondent and her social contact were of the same caste, and one to assess whether they had the same level of education.

Statistical Analyses

We first conducted bivariate analyses to test associations with ego latrine ownership using simple logistic regression and χ^2 tests. Variables that were at least marginally significant in the bivariate analyses ($p < 0.10$) were included in a multivariate model. We then interacted significant ($p < 0.05$) individual level covariates and the two homophily measures with the main predictor (alter's latrine ownership). To account for village level factors that may contribute to possible latrine ownership, such as geography or exposure to government latrine building campaigns, we included village level fixed effects in all of our multivariate models.

To account for autocorrelation that may arise from multiple observations of the same household, we used generalized estimating equation (GEE) procedures that cluster on household and assume an independent working correlation structure for the clusters. The GEE regression models in the tables provide parameter estimates in the form of beta coefficients, whereas the results reported in the text are in the form of odds ratios. We include many controls in these models, but the key coefficient is friend's latrine ownership (see Table 3.2).

Results

Summary statistics for the sample population are presented in Table 1. The mean age was 39, with approximately 54% of participants female. Roughly 31% of the sample was scheduled caste (25%) and scheduled tribe (6%), groups who have historically been the most disadvantaged and marginalized. The great majority of the participants were Hindu. Thirty-eight percent of the participants had no education, with the mean number of years of education being approximately 5 (SD 4.64). The mean number of participants per village was 221 (SD 70). There were a total of 117390 dyadic observations used in the analyses with a mean degree centrality score of 15.9 (SD 5.99). Some 30% of homes had a latrine, which is higher than the 17% average in rural India as a whole. The majority of the homes had private electricity (66%), with 29% using government electricity and 6% having no electricity at all. Figure 3.1 shows the network of a random village. Table 1 shows the proportion of individuals in each demographic group having a latrine. Latrine ownership is higher among those with higher caste status, more education, private electricity, and more technologically advanced roof construction (see Figure 3.2).

While the results of the bivariate analysis (see Table 3.2) suggest that almost all of the important predictors are statistically significant, when added together in a

multivariate analysis, the effects of language, religion, and all of the rooftype categories save RCC, diminish to non-significance. Consistent with prior research, those of the highest castes have the highest probability of owning a latrine versus those of the scheduled caste (OR 1.83, CI 1.41-2.38 for OBC and OR 2.90, CI 1.98-4.25 for General Caste). For each increase in education category, the odds of owning a latrine are 1.07 (CI 1.05-1.08) times that of the lowest education level. Those considered above the poverty level have an odds of latrine ownership 2.24 times higher (CI 1.81-2.78) than those considered below the poverty level. Household level predictors of latrine ownership are having private electricity versus having none (OR 3.28, CI 1.65-6.54)), RCC rooftype versus thatch (OR 5.26, CI 1.73-15.9), more rooms in the house (OR 1.40, CI 1.27-1.55), and more beds (OR 1.24, CI 1.10-1.38).

In the bivariate analysis there is a strong correlation between ego's latrine ownership with alter's latrine ownership. The odds of ego owning a latrine if alter does are 2.44 (CI 2.25-2.64)) times the odds of ego owning a latrine if alter does not. While this effect decreases in the multivariate model (OR 1.27 CI 1.16-1.39), it remains significant even when accounting for caste, education, religion, language, poverty level, household level income indicators, and village level fixed effects.

To see whether there are factors that influence the strength of the association between ego and alter latrine ownership, we included interactions in the model between alter's latrine ownership and other variables. Interactions with ego's caste and education level were not significant, suggesting that these variables do not influence the effect alter may have on ego (SA table SA3.1). However, the interaction tests for both of the homophily measures were significant. If ego and alter were the same caste, or had the same level of education then the association between ego's and alter's latrine ownership was slightly but significantly greater (Table 3.2). Our

measure of tie strength also significantly increases the association between ego and alter's latrine ownership (see Table 3.2). Although this is not a causal test, it is consistent with other work that suggests social contacts with strong ties are more likely to influence one another than those with weak ties [58]

Degree centrality showed a strong correlation with ego latrine ownership in the bivariate model, but in the multivariate model this association disappeared (see Table 3.3). In contrast, betweenness centrality and closeness centrality remained significant even when including numerous controls (Table 3.3). Each standard deviation increase in centrality is associated with increased odds of owning a latrine of 1.27 (95% CI 1.21-1.34) for betweenness, and 1.59 (95%CI 1.46-1.62) for closeness.

We also found significant interaction effects for both centrality measures with alter's latrine ownership. While those with the highest centrality measures were most likely to own a latrine, higher centrality measures also diminished the association with alter's latrine ownership. To better understand this dynamic, we stratified the data and repeated the analyses for those who were at the highest and lowest quartiles of both measures. We find that alter having a latrine increases the odds of individual latrine ownership by 1.40 (95% CI 1.23-1.61) for those with the lowest betweenness centrality scores versus an odds of 1.13 (95% CI 1.04-1.22) for those with the highest betweenness centrality. This is even more pronounced for closeness centrality. For those with a low closeness centrality score alter having a latrine increases the odds of ego having one by 1.72 (95% CI 1.47-2.01) versus an odds of 0.96 (95% CI 0.87-1.06) for those of high closeness centrality (Figure 3.3 and SA Tables SA3.3 and SA3.4). If a causal process underlies these associations, it suggests that people at

the center of the network are less likely to be influenced by their social contacts than people at the periphery.

Discussion

This study is the first to evaluate the association between latrine ownership and social network characteristics in a rural developing world setting. Using sociocentric data from 75 villages in rural India we tested whether the latrine ownership of a person's social contacts significantly predicted their own latrine ownership. We also explored whether the social network characteristics of an individual increased the probability of owning a latrine.

Controlling for caste, education, income, and village level fixed effects, we find that a person is significantly more likely to own a latrine if their social contacts also own latrines. We also find that homophily may help to explain some of this association: similarity between ego and alter in caste and education increase the likelihood that they exhibit the same outcome (they both own a latrine or neither owns a latrine).

Adding network measures to the models yielded several novel results. First, we found that people with stronger ties exhibit greater similarity in latrine ownership. This is consistent with other work that suggests close friends influence each other more than weak friends [59] but it is also possible that people who are more similar (reflected by their ownership of latrines) are more likely to become close.

Second, we found that network-level indicators of connectivity (as indicated by two separate measures of centrality, betweenness and closeness) were strong predictors of latrine ownership, even net of numerous controls for socioeconomic status. These controls also included a simple count of social contacts (degree centrality), suggesting that it is not just the social activity or popularity of an individual

that matters most. Instead, it is the individual's structural relationship to the network as whole.

Third, while those who are most central are most likely to own a latrine, they are also less likely to be influenced by their social contacts than those at the periphery (Figure 3). This raises an interesting dilemma for efforts to promote adoption: should we target people who are more influential at the center of the network or people who are more influenceable at the periphery[60]?

There are several possible explanations for these results. Government-sponsored latrine building campaigns were implemented in Karnataka during the 2000's though not all areas were uniformly covered [22]. However, the rate of household latrine ownership in our overall sample is consistent with that achieved after the government's sanitation campaign, suggesting that government-sponsored toilet building interventions took place in these villages. The high rate of latrine ownership among the most central individuals may thus be explained by the tendency for those most central in a community to receive the most exposure to outside interventions. This is both because their social position naturally offers them greater exposure to innovations spreading in the network, and also because many interventions attempt to target the most central individuals believing them to be socially influential [43 61 62].

People at the center of the network who have not yet adopted may thus be the least influenceable, which is consistent with previous research showing that central actors and those in denser networks can often be the most constrained by prevalent norms [9 35 45]. If latrine adoption has not become normative amongst the majority of the community (which is probable given that only 30% of households in this sample have

latrines) then those non-latrine owning central individuals may be the most difficult to persuade.

Thus, contrary to common assumptions about the role of centrality in the adoption of innovations, our results suggest that latrine-building interventions targeting those more peripheral to the network may be surprisingly efficacious. Not only are those more peripheral less likely to own a latrine, but they are also more likely to have friends without latrines. Most importantly, intervention efforts aimed at the more peripheral community members will potentially have larger multiplier effects, as the correlations between ego and alter latrine ownership are much higher in these groups. It is possible that in this context, previous intervention attempts were successful at reaching the more innovative central individuals but that the utility of outreach to the most central has its limitations. If a core group of central individuals are “hold-outs”, those more peripheral may be more amenable to change and as well as more likely to influence their less centrally located peers.

Qualitative studies assessing community-led sanitation programs have pointed to the importance of social norms in the success of programs aiming to increase latrine ownership and usage around the world. For the most part educational health campaigns and government subsidies have failed due to the inability of such approaches to address the social factors involved in latrine adoption, and have focused on convincing and educating *individuals* rather than *groups*. While we cannot definitively conclude that our results reflect the role of normative dynamics, they are suggestive of the possibility. While government-sponsored latrine building programs were unsuccessful at creating normative change, a minority of the population adopted latrines, some of them perhaps due to a social learning dynamic, having observed the benefits accrued to a friend or relation who did so. Overall

societal norms, however, did not shift to create the sort of social influence required to tip the majority into adopting latrine building.

Limitations of this study

The data used for this analysis was observational and cross-sectional so it is important to be cautious about inferring causation. However, because associations remain significant even when we have accounted for the major predictors of latrine ownership such as caste and education, future research using longitudinal data and experimental designs to investigate possible causality is certainly warranted.

Due to the static nature of the data, it is also difficult to control for homophily or the possibility that egos and alters are connected due to similarity in the kinds of characteristics that predict latrine ownership. We find some evidence that homophily may moderate the strength of association in latrine ownership between ego and alter, but it is not possible to determine whether this is because similar people choose to socialize with one another, or if it is because similar people are more liable to be influenced by one [63]. Ideally we would like to control for similarities at an initial point in time and then assess the correlations in latrine ownership over the course of a social relationship.

A further limitation is the lack of village level measures, meaning we cannot identify which environmental factors may be driving some of the results. Higher rates of latrine ownership were predicted in Benin among larger villages and those closer to urban centers [46], and in Orissa among those who had been exposed to higher quality government latrine campaigns [23]. However, by controlling for village level fixed effects, we have been able to statistically account for those unmeasured predictors, so they should not affect the estimates of individual and household level effects that we present here.

Finally, this data does not measure latrine usage. While latrine ownership is certainly an important step towards actually using a latrine, it is not a guarantee. Work with CLTS highlighted the fact that it takes more than actually owning a latrine to ensure that people are really using them [26]. Again, future research can investigate the correlation between latrine ownership, social network factors, and actual latrine usage.

Conclusion

Despite the limitations of this study, it has several important strengths. This is the first study to quantitatively characterize the social network characteristics associated with latrine ownership. While an enormous amount of time and energy is being devoted to promoting latrine building and usage in South Asia, little quantitative evidence exists to inform programs and policies. The large number of villages (75) represented in our study provides results that may be more generalizable than those based upon only one or two. While we cannot infer causal effects with this data, our results suggest that social dynamics are an integral aspect of latrine adoption and that these dynamics require a more complex approach than simply identifying the most central individuals and appealing to them. Qualitative studies have suggested that for latrine building to succeed it must become normative within a community rather than a matter of individual preference. Our results are consistent with these conclusions.

Table 3.1: Summary Statistics Individual Survey from 75 villages in rural Karnataka India N=16605

	Mean (SD)	
Age in years	39.30 (12.5)	
Education mean years	5.02 (4.65)	
Number of rooms in house	2.62 (1.54)	
Number of beds in house	1.05 (1.62)	
Tie strength[$]	3.12 (2.45)	
Total degree centrality	15.90 (5.99)	
Betweenness centrality[#]	7.60 (0.92)	
Closeness centrality[*]	29.60(3.19)	
	Proportion of total sample	**Proportion with household latrine**
Gender (% female)	55.10%	30.00%
Caste		
Scheduled caste	25.27%	12.50%
Scheduled tribe	6.05%	12.80%
OBC	56.45%	37.20%
General	12.23%	44.70%
Ration card category		
Above poverty	17.40%	55.70%
Below poverty	66.90%	22.10%
Does not own card	15.70%	38.10%
Religion		
Hindu	96.11%	29.80%
Islamic	3.79%	46.70%
Language		
Kannada	74.92%	31.70%
Tamil	3.98%	23.30%
Telugu	17.33%	22.80%
Hindi	3.59%	48.50%
Other	.181%	43.30%
Roof-type		
Thatch	1.98%	6.69%
Tile	31.38%	17.06%
Stone	30.55%	32.57%
Sheet	17.96%	27.89%
RCC	13.74%	67.47%
Other	4.11%	15.21%
Household electricity		
Private	65.53%	41.38%
Govt.	28.65%	10.32%
No	5.82%	5.67%
Latrine ownership yes	30.4%	

$ tie strength value calculated using full dyadic dataset (n=117390)
#betweenness (be) centrality transformation: log(be+1)
centrality (tc) transformation: tc 100

Table 3.2. Results of GEE logistic regression using dyadic observations on predictors of latrine ownership in rural Karnataka India N=117390

	Bivariate analyses	Multivariate Model #1 main effect	Multivariate Model #2 Tie strength interaction	Multivariate Model # 3 Same caste binary interaction	Multivariate Model #4 Same education binary interaction
	Beta (SE)	Beta (SE)	Beta (SE)	Beta (SE)	Beta (SE)
Alter's latrine ownership	**0.891*** (0.032)**	**0.178*** (0.034)**	**0.099* (0.04)**	**0.042 (0.061)**	**0.139*** (0.035)**
Education in years	0.126*** (0.005)	0.062*** (0.006)	0.061*** (0.006)	0.064*** (0.007)	0.06*** (0.007)
Number of rooms in house	0.611*** (0.038)	0.344*** (0.046)	0.342*** (0.046)	0.341*** (0.046)	0.342*** (0.046)
Number of beds in house	0.588*** (0.06)	0.189*** (0.054)	0.189*** (0.054)	0.189*** (0.054)	0.189*** (0.054)
Ration card					
Above poverty line	ref	ref	ref	ref	ref
Below poverty line	-1.512*** (0.078)	-0.749*** (0.098)	-0.744*** (0.098)	-0.745*** (0.098)	-0.745*** (0.098)
No ration card	-0.663*** (0.098)	-0.353** (0.127)	-0.347** (0.127)	-0.348** (0.127)	-0.348** (0.127)
Caste					
Scheduled caste	ref	ref	ref	ref	ref
Scheduled tribe	-0.113 (0.197)	0.141 (0.238)	0.15 (0.239)	0.123 (0.24)	0.149 (0.239)
OBC	1.378*** (0.097)	0.66*** (0.115)	0.717*** (0.113)	0.709*** (0.113)	0.716*** (0.113)
General	1.795*** (0.12)	1.084*** (0.17)	1.12*** (0.168)	1.091*** (0.168)	1.12 ***(0.168)
Household electricity					
Private	ref	ref	ref	ref	ref
Govt	-1.842*** (0.098)	-0.956*** (0.113)	-0.949*** (0.112)	-0.949*** (0.112)	-0.95*** (0.112)
No	-2.391*** (0.334)	-1.438*** (0.298)	-1.434*** (0.297)	-1.436*** (0.297)	-1.433*** (0.297)
Roof-type					
Thatch	ref	ref	ref	ref	ref
Tile	1.131* (0.444)	-0.284 (0.497)	-0.267 (0.497)	-0.269 (0.497)	-0.267 (0.497)
Stone	2.02*** (0.443)	0.563 (0.503)	0.568 (0.502)	0.565 (0.503)	0.568 (0.502)
Sheet	1.754*** (0.446)	0.486 (0.498)	0.504 (0.497)	0.501 (0.498)	0.504 (0.497)
RCC	3.399*** (0.447)	1.428* (0.512)	1.433** (0.511)	1.432*** (0.512)	1.434*** (0.512)
Other	0.746 (0.488)	-0.49 (0.54)	-0.48 (0.54)	-0.479 (0.54)	-0.48 (0.54)
Religion					
Hindu	ref	ref	ref	ref	ref

Table 3.2 continued

	Bivariate analyses	Multivariate Model #1 main effect	Multivariate Model #2 Tie strength interaction	Multivariate Model # 3 Same caste binary interaction	Multivariate Model #4 Same education binary interaction
Islamic	0.819***	-0.132			
	(0.146)	(0.487)			
Language					
Kannada	ref				
Tamil	-0.403*	-0.213			
	(0.17)	(0.212)			
Telugu	-0.514***	-0.022			
	(0.086)	(0.118)			
Hindi	0.769***	0.69			
	(0.151)	(0.508)			
Tie Strength			**-0.009**		
			(0.006)		
Tie Strength*alter latrine			**0.026****		
			(0.009)		
Same caste binary				**-0.156****	
				(0.053)	
Same caste* alter latrine				**0.183***	
				(0.072)	
Same education binary					**-0.086***
					(0.04)
Same education * alter latrine					**0.219*****
					(0.051)
Multivariate models include village level fixed effects not shown					

Table 3.3. Results of a GEE logistic regression analysis using dyadic observations testing the effect of an individual's network characteristics on the probability of latrine ownership, rural Karnataka India N=117390

	Bivariate[%]	Multivariate[@]	Multivariate controlling for total degree	Interaction with alter latrine* betweenness	Interaction with alter latrine* closeness
	Beta (SE)	Beta (SE)	Beta (SE)	Beta (SE)	Beta (SE)
Betweenness	0.263*** (0.027)	0.116*** (0.034)	0.105* (0.046)	0.141*** (0.050)	
Closeness	0.145*** (0.013)	0.051** (0.017)	0.055* (0.023)		0.071*** (0.018)
Total Degree	0.023*** (0.003)	0.006 (0.004)			
Betweenness* alter latrine				-0.070*** (0.037)	
Alter latrine				0.756*** (0.301)	
Closeness* alter latrine					-0.060*** (0.011)
Alter latrine					2.003*** (0.342)

%All univariate models were run including village level fixed effects

@All multivariate models were run including respondent caste, education, ration card category, household electricity, household rooftype, number of rooms in house, number of beds in house, alter latrine ownership, and village level fixed effects

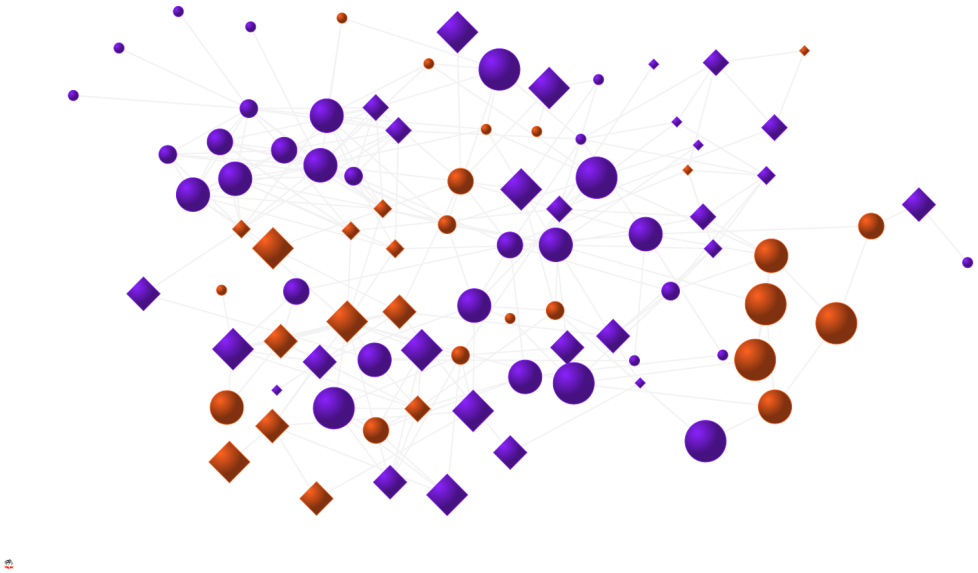

Figure 3.1: A network depiction of a randomly chosen village from the dataset. Orange nodes are latrine owning and purple nodes are latrine non-owning. Circles are higher caste (OBC and general), diamonds are lower castes (scheduled tribe and scheduled caste). The size of the node increases according to the proportion of the individual's friends that own latrines. Note that: 1) latrine owners are predominantly higher caste (OBC or general) shown by the orange circles 2) latrine owners tend to be clustered together as evident by the groupings of orange versus the groupings of purple 3) a higher proportion of latrine owners have friends with latrines shown by the high proportion of larger orange nodes 4) those latrine owners who have a smaller proportion of friends with latrines tend to be more central to the network evident by the smaller orange nodes near the center versus the smaller purple nodes near the periphery.

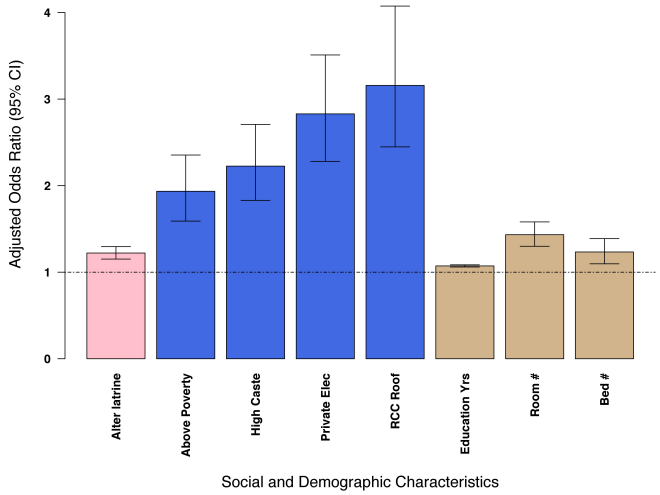

Figure 3.2. The adjusted odds ratio of latrine ownership by various social and demographic characteristics (95% CI). The full model was simplified so that the blue bars depict binary predictors of those factor categories most predictive of latrine ownership. The tan bars are continuous variables with odds ratios calculated at a 1 unit increase. All variables adjusted for all other variables depicted in the figure plus village level fixed effects (not shown). Simplified model shown in SA table SA3.5.

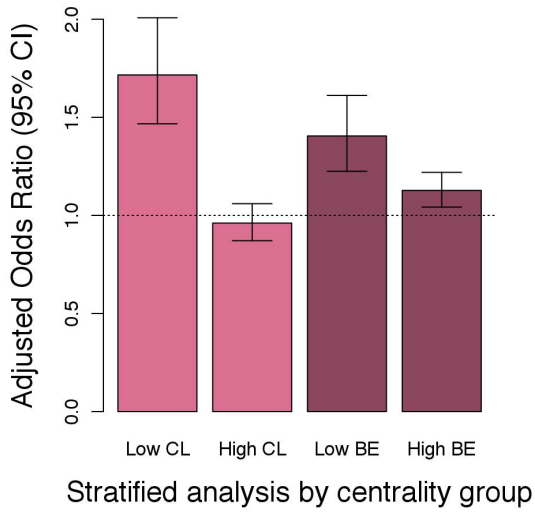

Figure 3.3. Results of a stratified analyses of the impact of alter's latrine ownership on ego's latrine ownership by ego low (1st quartile) and high (fourth quartile) closeness centrality (CL) and betweenness centrality (BE). Here we see that the impact of alter's latrine ownership on ego's latrine ownership differs according to ego's centrality. Those more peripheral to the network are potentially more susceptible to the social effect of alter's latrine ownership than those more central. All models run controlling for individual and household covariates plus village level fixed effects. Full models shown in SA tables SA3.3 and SA3.4.

References

1. Black RE, Cousens S, Johnson HL, Lawn JE, Rudan I, Bassani DG, et al. Global, regional, and national causes of child mortality in 2008: a systematic analysis. *The Lancet* 2010;375(9730):1969-87.

2. Pruss A, Kay D, Fewtrell L, Bartram J. Estimating the burden of disease from water, sanitation, and hygiene at a global level. *Environmental health perspectives* 2002;110(5):537-42.

3. Pattanayak SK, Yang JC, Dickinson KL, Poulos C, Patil SR, Mallick RK, et al. Shame or subsidy revisited: social mobilization for sanitation in Orissa, India. *Bulletin of the World Health Organization* 2009;87(8):580-87

4. Bonu S, Kim H. Sanitation in India: Progress, Differentials, Correlates, and Challenges. In: Bank AD, editor. *South Asia Occasional Paper Series No. 2*. Mandaluyong City: Asian Development Bank, 2009.

5. Nath KJ. Home hygiene and environmental sanitation: a country situation analysis for India. *International Journal of Environmental Health Research* 2003;13(S1):19-28 %@ 0960-3123.

6. Kumar S, Vollmer S. Does improved sanitation reduce diarrhea in children in rural India? *Health Economics* 2011;under revision.

7. WHO, UNICEF. Meeting the MGD drinking water and sanitation target: the urban and rural challenge of the decade. In: WHO, UNICEF, editors. Switzerland, 2006.

8. Baligar SVP. Human Development Report Karnataka 2005 In: http://planningcommission.nic.in/plans/stateplan/sdr_pdf/shdr_kar05.pdf, editor: Planning and Statistics Department Government of Karnataka, 2006.

9. Meenakshisundaram S. Feeling the Pulse: A Study of the Total Sanitation Campaign in Five States. New Delhi: WaterAid India, 2008.

10. Dickinson K, Pattanayak S. Open Sky Latrines: Do social effects influence technology adoption in the case of a (very) impure public good? . *Working Paper, Duke University* 2012.

11. Veerashekharappa, Bhide S. Sanitation Strategies in Karnataka: A Review: Institute for Social and Economic Change, 2009.

12. Jenkins MW, Curtis V. Achieving the [] good life': Why some people want latrines in rural Benin. *Social Science & Medicine* 2005;61(11):2446-59 %@ 0277-9536.

13. Dyalchand A, Khale M, Vasudevan S, Kale N. What Communication and Institutional Arrangements Influence Sanitation Related Social Norms in Rural

India? In: L M, S M, editors. *Shit Matters: The Potential of Community-Led Total Sanitation* London: Practical Action, 2008.

14. Chambers R. Going to Scale with Community-Led Total Sanitation: Reflections on Experience, Issues and Ways Forward. *IDS Practice Papers* 2009;2009(1):01-50 %@ 2040-0225.

15. Cialdini RB, Kallgren CA, Reno RR. A focus theory of normative conduct: A theoretical refinement and reevaluation of the role of norms in human behavior. *Advances in experimental social psychology* 1991;24(20):1-243.

16. Lapinski MK, Rimal RN. An Explication of Social Norms. *Communication Theory* 2005;15(2):127-47.

17. Sunstein CR. Social norms and social roles. *Columbia Law Review* 1996;96(4):903-68 %@ 0010-1958.

18. Burke MA, Young HP. Social Norms. In: Bisin Alberto BJ, Jackson Matthew, editor. *Handbook of Social Economics*. Amsterdam: North Holland, 2009.

19. Bwire B. Breaking shit taboos: CLTS in Kenya. In: Development TIIfEa, editor. *Participatory Learning and Action*, 2010:91-96 %@ 1357-938X.

20. Kohler HP, Behrman JR, Watkins SC. The density of social networks and fertility decisions: Evidence from South Nyanza District, Kenya. *Demography* 2001;38(1):43-58 %@ 0070-3370.

21. Montgomery MR, Casterline JB. Social learning, social influence, and new models of fertility. *Population and Development Review* 1996;22:151-75 %@ 0098-7921.

22. Rosenquist J, Murabito J, Fowler JH, Christakis NA. The spread of alcohol consumption behavior in a large social network. *Annals of Internal Medicine* 2010;152(7):426.

23. Christakis NA, Fowler JH. The spread of obesity in a large social network over 32 years. *New England journal of medicine* 2007;357(4):370-79 %@ 0028-4793.

24. Christakis NA, Fowler JH. Social Contagion Theory: Examining Dynamic Social Networks and Human Behavior. *Statistics in Medicine* 2012;in press.

25. Shakya H, Christakis NA, Fowler JH. Parental network influences on adolescent outcomes. *Archives of Pediatrics and Adolescent Medicine* 2012;Forthcoming.

26. Gayen K, Raeside R. Social networks and contraception practice of women in rural Bangladesh. *Social Science & Medicine* 2010;71(9):1584-92.

27. Rothenberg RB, Potterat JJ, Woodhouse DE, Darrow WW, Muth SQ, Klovdahl AS. Choosing a centrality measure: epidemiologic correlates in the Colorado Springs study of social networks. *Social Networks* 1995;17(3-4):273-97.

28. Christley RM, Pinchbeck GL, Bowers RG, Clancy D, French NP, Bennett R, et al. Infection in social networks: using network analysis to identify high-risk individuals. *Am J Epidemiol* 2005;162(10):1024-31.

29. Christakis NA, Fowler JH. Social network sensors for early detection of contagious outbreaks. *PloS one* 2010;5(9):e12948 %@ 1932-6203.

30. Valente T. *Network models and methods for studying the diffusion of innovations*, 2005.

31. Fujimoto K, Valente TW, Pentz MA. Network structural influences on the adoption of evidence-based prevention in communities. *Journal of Community Psychology* 2009;37(7):830-45 %@ 1520-6629.

32. Banerjee A, Chandrasekhar AG, Duflo E, Jackson MO. The Diffusion of Microfinance: Mimeo. http://econ-www. mit. edu/files/6910, 2011.

33. Gayen K, Raeside R. Communicative actions, women's degree of social connectedness and child mortality in rural Bangladesh. *Child: Care, Health and Development* 2010;36(6):827-34 %@ 1365-2214.

34. Jenkins MW, Cairncross S. Modelling latrine diffusion in Benin: towards a community typology of demand for improved sanitation in developing countries. *Journal of water and health* 2010;8(1):166-83 %@ 1477-8920.

35. Valente TW. Network models and methods for studying the diffusion of innovations. *Models and methods in social network analysis* 2005:98-116.

36. Jackson MO, Rodriguez-Barraquer T, Tan X. Social patterns and social quilts: network patterns of favor exchange. *American Economic Review* 2012.

37. Marin A, Wellman B. Social network analysis: An introduction. *The SAGE Handbook of Social Network Analysis* 2011:11.

38. Morris SS, Carletto C, Hoddinott J, Christiaensen LJM. Validity of rapid estimates of household wealth and income for health surveys in rural Africa. *Journal of epidemiology and community health* 2000;54(5):381-87 %@ 1470-2738.

39. Marin A. Are respondents more likely to list alters with certain characteristics?Implications for name generator data. *Social Networks* 2004.

40. Marin A, Hampton KN. Simplifying the Personal Network Name Generator. *Field Methods* 2007;19(2):163.

41. Proctor C, Loomis C. Analysis of Sociometric Data. In: Holland P, editor. *Research Methods in Social Relations*. New York: Dryden Press, 1951.

42. Sabidussi G. The centrality index of a graph. *Psychometrika* 1966;31(4):581-603.

43. Freeman LC. A set of measures of centrality based on betweenness. *Sociometry* 1977:35-41.

44. Wasserman S, Faust K. *Social Network Analysis: Methods and Applications*. Cambridge: Cambridge University Press, 1994.

45. McPherson M, Smith-Lovin L, Cook JM. Birds of a feather: Homophily in social networks. *Annual Review of Sociology* 2001:415-44 %@ 0360-572.

46. Bond R, Fariss C, Jones F, Settle A, Marlow C, Fowler JH. A massive scale experiment in social influence and political mobilization.*?* 2012.

47. Brown JJ, Reingen PH. Social ties and word-of-mouth referral behavior. *Journal of Consumer Research* 1987:350-62 %@ 0093-5301.

48. Christakis NA, Fowler JH. Contagion in prescribing behavior among networks of doctors. *Marketing Science* 2010;30(2):213-16.

49. Borgatti SP, Mehra A, Brass DJ, Labianca G. Network analysis in the social sciences. *science* 2009;323(5916):892-95 %@ 0036-8075.

50. Valente T. *Social networks and health: Models, methods, and applications*: Oxford Univ Pr, 2010.

51. Centola D. The spread of behavior in an online social network experiment. *science* 2010;329(5996):1194-97 %@ 0036-8075.

Supplementary Appendix

Table SA3.1: Results of GEE logistic regression using dyadic observations on predictors of latrine ownership in rural Karnataka India: interactions N=117390

	Ego caste*alter latrine interaction			Ego education*alter latrine interaction		
	Beta	SE	P	Beta	SE	P
Alter's latrine ownership	0.25	0.11	0.02	0.21	0.08	0.01
Education in years	0.07	0.01	0.00	0.06	0.01	0.00
Number of rooms in house	0.39	0.06	0.00	0.39	0.06	0.00
Number of beds in house	0.19	0.06	0.00	0.19	0.06	0.00
Ration card						
Above poverty line	ref			ref		
Below poverty line	-0.78	0.15	0.00	-0.78	0.15	0.00
No ration card	-0.22	0.22	0.31	-0.23	0.22	0.30
Caste						
Scheduled caste	ref			ref		
Scheduled tribe	0.36	0.36	0.31	0.24	0.34	0.49
OBC	0.78	0.17	0.00	0.73	0.17	0.00
General	1.20	0.27	0.00	1.17	0.26	0.00
Household electricity						
Private	ref			ref		
Govt	-1.17	0.17	0.00	-1.16	0.17	0.00
No	-1.62	0.50	0.00	-1.61	0.50	0.00
Roof-type						
Thatch	-0.23	0.75	0.76	ref		
Tile	0.46	0.76	0.54	-0.23	0.75	0.75
Stone	0.43	0.75	0.56	0.46	0.75	0.54
Sheet	1.30	0.77	0.09	0.43	0.75	0.57
RCC	-0.51	0.81	0.53	1.29	0.76	0.09
Other	-0.23	0.75	0.76	-0.51	0.81	0.53
Caste * Alter latrine						
Scheduled caste	ref					
Scheduled tribe	**-0.43**	**0.25**	**0.09**			
OBC	**-0.16**	**0.12**	**0.20**			
General	**-0.10**	**0.16**	**0.54**			
Education* Alter latrine				**-0.01**	**0.01**	**0.13**

**Controls for village level fixed effects not shown*

Table SA3.2. Results of GEE logistic regression using dyadic observations on predictors of latrine ownership in rural Karnataka India*: including alter level controls N=117390

	Beta	SE	P
Alter's latrine ownership	***0.12***	***0.04***	***0.00***
Education in years	0.06	0.01	< 2e-16
Number of rooms in house	0.34	0.05	0.00
Number of beds in house	0.19	0.05	0.00
Ration card			
Above poverty line			
Below poverty lin	-0.74	0.10	0.00
No ration card	-0.34	0.13	0.01
Caste			
Scheduled caste			
Scheduled tribe	0.20	0.24	0.41
OBC	0.75	0.12	0.00
General	1.14	0.17	0.00
Household electricity			
Private			
Govt	-0.95	0.11	< 2e-16
No	-1.43	0.30	0.00
Roof-type			
Thatch			
Tile	-0.27	0.50	0.59
Stone	0.57	0.50	0.26
Sheet	0.50	0.50	0.31
RCC	1.43	0.51	0.01
Other	-0.48	0.54	0.38
ALTERS ATTRIBUTES:			
Education in years	0.01	0.00	0.05
Number of rooms in house	0.00	0.01	0.78
Number of beds in house	0.02	0.01	0.07
Ration card			
Above poverty line			
Below poverty line	-0.05	0.04	0.18
No ration card	-0.07	0.05	0.14
Caste			
Scheduled caste			
Scheduled tribe	-0.10	0.10	0.30
OBC	-0.07	0.06	0.28
General	-0.08	0.08	0.34
Household electricity			
Private			
Govt	0.02	0.04	0.69
No	-0.04	0.07	0.60
Roof-type			
Thatch	-0.21	0.13	0.10
Tile	-0.12	0.13	0.38
Stone	-0.19	0.13	0.14
Sheet	-0.10	0.13	0.47
RCC	-0.21	0.15	0.18
Other	-0.21	0.13	0.10

**Controls for village level fixed effects not shown*

Table SA3.3. Results of GEE logistic regression using dyadic observations on predictors of latrine ownership in rural Karnataka India N=117390
with separate models for data stratified by closeness centrality[%] (figure 3)

	Low closeness centrality (<27.4) N=17738			**High closeness centrality (>31.7)* N=43907**		
	Beta	SE	P	Beta	SE	P
Alter's latrine ownership	**0.54**	**0.08**	**<.001**	**-0.04**	**0.05**	**0.38**
Education in years	0.06	0.01	0.00	0.06	0.01	0.00
Number of rooms in house	0.41	0.08	0.00	0.30	0.07	0.00
Number of beds in house	0.35	0.09	0.00	0.11	0.08	0.14
Ration card						
Above poverty line						
Below poverty line	-0.27	0.20	0.16	-0.82	0.17	0.00
No ration card	0.05	0.21	0.81	-0.34	0.23	0.14
Caste						
Scheduled caste						
Scheduled tribe	0.11	0.39	0.77	0.25	0.41	0.55
OBC	0.88	0.20	0.00	0.85	0.22	0.00
General	0.90	0.30	0.00	1.02	0.28	0.00
Household electricity						
Private						
Govt	-0.67	0.18	0.00	-1.21	0.22	0.00
No	-2.37	0.60	0.00	-1.35	0.47	0.00
Roof-type						
Thatch				-38.12	0.38	< 2e-16
Tile	-1.95	0.83	0.02			
Stone	-1.04	0.80	0.19	1.02	0.27	0.00
Sheet	-1.23	0.81	0.13	0.84	0.24	0.00
RCC	-0.14	0.83	0.86	1.71	0.30	0.00
Other	-1.87	0.86	0.03	-0.52	0.47	0.27

% all models adjusted for village level fixed effects not shown
**because of small cell size of thatch roof in this model, we used tile as the reference instead*

Table SA3.4. Results of GEE logistic regression using dyadic observations on predictors of latrine ownership in rural Karnataka India N=117390 with separate models for data stratified by betweenness centrality[%] (figure 3)

	Low betweenness centrality (<7.07) N=19405			**High betweenness centrality (>8.20) N=41916**		
	Beta	SE	P	Beta	SE	P
Alter's latrine ownership	**0.34**	**0.07**	**0.00**	**0.12**	**0.04**	**0.01**
Education in years	0.05	0.01	0.00	0.07	0.01	0.00
Number of rooms in house	0.36	0.07	0.00	0.39	0.06	0.00
Number of beds in house	0.17	0.10	0.10	0.19	0.06	0.00
Ration card						
Above poverty line						
Below poverty line	-0.61	0.18	0.00	-0.78	0.15	0.00
No ration card	-0.34	0.21	0.10	-0.22	0.22	0.31
Caste						
Scheduled caste						
Scheduled tribe	-0.14	0.41	0.73	0.24	0.34	0.41
OBC	0.62	0.20	0.00	0.73	0.17	0.00
General	0.77	0.28	0.01	1.18	0.26	0.00
Household electricity						
Private						
Govt	-0.92	0.18	0.00	-1.17	0.17	0.00
No	-1.53	0.45	0.00	-1.62	0.50	0.00
Roof-type						
Thatch						
Tile	-0.34	0.82	0.67	-0.23	0.75	0.75
Stone	0.81	0.82	0.33	0.46	0.75	0.55
Sheet	0.69	0.82	0.39	0.43	0.75	0.57
RCC	1.68	0.84	0.05	1.29	0.76	0.09
Other	0.10	0.89	0.91	-0.51	0.81	0.53

% all models adjusted for village level fixed effects not shown

Table SA3.5: Simplified model using binary predictors[*] (figure 2)

	Beta	SE	P
Alter latrine	0.20	0.03	0.00
High caste	0.81	0.10	0.00
Education in years	0.07	0.01	< 2e-16
RCC roof	1.16	0.13	< 2e-16
# of beds	0.21	0.06	0.00
# of rooms	0.36	0.05	0.00
Private electricity	1.04	0.11	< 2e-16
Above poverty line	0.66	0.10	0.00

**Controls for village level fixed effects not shown*

Chapter 4

New perspectives on latrine ownership in rural India: the effect of social network communities

Abstract

Background: Social norms play a significant role in the adoption of important health innovations, including household latrine ownership in developing countries. A key challenge in changing social norms is to identify the social groups to whom individuals turn regarding normative expectations.

Methods: This study tests social network correlates of latrine ownership in rural India at three distinct level of social interaction: the individual, the social network community, and the village. Participants from 75 villages in rural India provided the names of their social contacts as well as their own relevant demographic and household characteristics. Using a social network community detection algorithm we derived 392 mutually exclusive communities from within the overall network. We then created community and village level aggregate measures of possible predictors of latrine ownership, including group level latrine ownership proportions, and tested them against individual latrine ownership.

Results: Using a random effects model to account for within-group correlations, we found that the latrine ownership proportions of social contacts at all three levels of social interaction were positively and significantly correlated with individual latrine ownership. Of these, however, the effect of community level latrine ownership is the strongest suggesting the possible presence of social norms as a determinant of latrine ownership.

Conclusion: If the village unit is not the primary source of normative expectations for individuals, village level interventions aimed at changing norms may be inefficient. Interventions directed towards normatively determined behaviors may be best implemented using a social network approach to identify and engage socially relevant communities.

Background and significance

Millennium Development Goals were adopted by countries around the world as part of a united effort to improve the health and well being of the most world's most impoverished people[1]. As part of that worldwide initiative, Target 7.C aims to reduce by one half, the number of people in the world without access to basic sanitation. Poor sanitation, including the lack of clean functioning toilets, is a major factor contributing to morbidity and mortality from infectious disease [2 3]. Diarrheal diseases are one of the most common causes of death for children under the age of 5, with almost 50% of those deaths occurring in India, Nigeria, Pakistan, Democratic Republic of Congo, and China [4]. Although efforts to improve sanitation are underway, people who are from disadvantaged social groups, who are impoverished, or who live in rural areas are still unlikely to have access to basic sanitation. For those without access to basic sanitation, open defecation is the accepted alternative, and has been practiced in many communities for centuries[5]. These problems are particularly pronounced in South Asia, including India, where over half of the worldwide affected population resides [6].

Although the government of India has devoted considerable resources to latrine building campaigns across the country over the last decade, results have been mixed, with a considerable disparity still existing between those of higher socio-economic status versus those of lower[1 7]. The most successful latrine building campaigns seem to be those patterned after Community Led Total Sanitation (CLTS) which utilizes a norm-based approach to encourage villages to become open-defecation free. [8]

Social Norms

A social norms perspective on behavior change considers the choices of individuals to be significantly impacted by the opinions and choices of those in their salient *reference groups*[9 10], or the people to whom they look for boundaries and expectations regarding social behavior. Cialdini was the first to differentiate what he called *descriptive* and *injunctive* norms. Descriptive norms are those regular behaviors that can be observed within a community, and can be adopted regardless of the expectations of others.[11] Injunctive norms reflect community standards as to what is expected behavior, and are therefore enforced through social sanctions. Sanctions can be negative for transgressions or positive for conformity and can be as blatant as a public shaming or as subtle as unspoken disapproval[12-14]. Furthermore, sanctions can come from a third party and not necessarily from a person directly involved in the behavior in question[15]. Cialdini argues that descriptive and injunctive norms motivate behavior through distinct cognitive mechanisms [12 16]. Descriptive norms operate mainly through a need for efficiency. Research has shown that simply the observance of others' behaviors can motivate action, usually unbeknownst to the individual adopting the behavior[17]. Injunctive norms on the other hand are more complex, and may be conflicting depending on the situation, as their primary function is to satisfy the goal of maintaining social relationships[16]. Work by Jacobson and colleagues demonstrated that descriptive and injunctive norms trigger different cognitive mechanisms with descriptive norms priming associations of "accuracy and efficiency", and injunctive norms priming those same associations plus those of social approval [16]. Injunctive norms can be difficult to change due to the reciprocity of normative expectations: it can be very difficult for any one individual to defy a social norm even when the expected outcome may be personally beneficial because the

individual can only avoid sanctions if the entire community is willing to change simultaneously [18 19]. Programs such as CLTS operate with the goal of triggering normative change by creating injunctive norms against negative behaviors like open-defecation.

A key task in the goal of fostering normative change is to identify the valid reference group for each individual in the population of interest. Normative change hinges on the fact that the individuals involved place a value upon the expectations of those in that reference group, and furthermore that they believe in the legitimacy of those expectations [10 14]. Most research surveys have not collected data specific to social norms, requiring researchers to attempt to create measures of possible norms through the aggregation of existing measures and to infer reference groups from higher level clusters, such as DHS sampling clusters, villages, or counties which are already present in the data. Individual level data can be aggregated at the cluster level and used to predict individual outcomes, employing multilevel modeling to account for the clustering of standard errors across groups. From these aggregated data, it is possible to infer possible community level norms. This method of modeling normative influences has been used in a variety of studies spanning outcomes as diverse as the use of contraception[20 21], female genital cutting [22], domestic violence [23], youth aggression [24], alcohol abuse [25], and adolescent smoking [26]. The utility of the method, however, depends on how the higher-level community units are defined and measured.

Community detection

Social network data is typically collected by asking individuals with whom they share certain types of relationships, for instance with whom they discuss important matters, or to whom they would go for help in an emergency. The information from a

certain population of people is then aggregated to create a nexus of relationships within which each individual is embedded. While these large social networks can provide valuable information regarding the greater social context of an individual, they are not necessarily reflective of the more cohesive communities to whom the individual refers when making socially relevant decisions. It may be necessary therefore to identify subgroups within the larger network that can also provide information beyond direct friend-to friend-ties. How best to accomplish this goal has been at the heart of much network research over the last few decades[27-29]. Community detection methods identify sets of individuals that have relatively strong within group ties, and weak between group ties[27]. While it is possible to measure communities by directly asking respondents, this methods often creates communities that overlap, making it difficult to tease out important differences in between-group versus within-group characteristics[30]. Furthermore, this method is not objective and can compound the respondent bias already inherent in self-report measures. As an alternative, several network algorithms have been recently developed to mathematically define network communities using iterative calculations[27]. While no one method is decidedly superior, there are several that have been empirically validated as capable of detecting network sub-groups that are reflective of real-world social communities[27].

Despite the increasing interest in community detection, very few empirical studies have utilized network-based communities to help understand the impact of higher-level social dynamics on health behaviors and outcomes. This may be the result of the dearth of diverse network data, as those who have utilized these methods have mainly focused on the impact of adolescent peer networks on

substance abuse[30 31 32], a topic about which several high quality network datasets are available. Health behavior research, however is increasingly focused on higher level social forces in relation to many diverse outcomes, as is reflected in the large body of literature on social ecology, which focuses on the reciprocal interactions between the individual and the various layers of their social environment, including community at different levels [33 34]. Indeed, research on norms cannot progress without this level of analysis, as norms are generated, diffused, held, and ultimately shifted within communities.

This study

A large body of qualitative research points to the fact that, above and beyond the more obvious demographic factors, social dynamics can influence the decision to build a latrine [3 35]. One recent study found a significant correlation between the latrine ownership of a person's nominated social contacts and their own latrine ownership [36]. Results from a study assessing the effects of Community Led Total Sanitation (CLTS) demonstrated that galvanizing entire communities was a necessary aspect of a latrine adoption campaign, and that the adoption decisions of the entire community were one of the strongest influences in a household's decision to build a latrine [5]. The most successful latrine building campaigns have been those initiated by CLTS, that successfully shift the norms of the community towards intolerance of open defecation, and community willingness to invest in building latrines [8]. Yet government subsidies for building latrines have been largely unsuccessful[37]. Pattanayak and colleagues (2009) have found, in fact, that for those above the poverty line, social shaming is a more economical and efficacious strategy for promoting latrine adoption than the use

of subsidies, although monetary support may be necessary for those below the poverty line[3] .

Despite this preliminary evidence, there is very little quantitative research investigating predictors of latrine usage, and only one that utilized social network data to actually match respondents with their network members in order to test empirically whether network social effects predict latrine ownership decisions[36]. Given the tremendous public health impact of open defecation, and in response, the significant effort worldwide to promote latrine usage, rigorous quantitative research analyzing the determinants of latrine ownership is essential.

Using network data collected in rural India we test the relationship between social network predictors and individual latrine usage at 3 levels: that of an individual's direct social contacts, that of the village within which the individual lives, and that of his/her distinct network-defined communities. We also include higher-level aggregated measures of individual level characteristics to determine whether demographic characteristics known to predict latrine ownership such as caste or education are predictive of individual level latrine ownership at the community or village level. In other words, is a person more likely to own a latrine if they are surrounded by people of higher education or higher caste? Given the village level focus and impact of latrine building campaigns, we hypothesize that village level social effects will be the strongest, and that individual and community network effects will also contribute, but to a lesser degree.

Methods

Data

We utilize *sociocentric* network data collected by Abhijit Banerjee and colleagues from 75 villages in rural Karnataka, in Southern India [38 39]. Sociocentric

studies focus on a small population and attempt to ascertain all of the social relationships within a set of interconnected individuals[40]. This is in contrast to *egocentric* network studies that focus on a larger population and attempt to ascertain all of the social relationships of a set of randomly-chosen individuals that are usually not connected to one another. Whereas egocentric data may help to improve the representativeness of a sample for a large population, sociocentric data allows measurement of larger network structures (like communities) and individual level network measures based on them.

The Karnataka data was collected as part of a study to understand the network diffusion of micro-finance. A complete census was taken by interviewing one person within each household, including information regarding household characteristics such as latrine ownership and roof construction. Individual surveys were used to collect demographic and network data from adult women from the ages of 18-57 and their eligible spouses from approximately half of the eligible households stratified by religion and geographic sub-location [38].

Outcome, demographic, and household level measures

Participants reported their age, gender, religion (Hindu or Muslim), and mother tongue (Kannada, Tamil, Telugu, or Hindi). Participants were also asked to identify to which caste they belong (scheduled caste, scheduled tribe, obc (other backward caste), or general). Education was measured using 16 levels ranging from none to higher degree. Household quality variables included roof type (6 categories), number of rooms in the home, number of beds in the home, and household electricity (private electricity, government electricity, or no electricity). Consistent with prior work in traditional agrarian societies in which data regarding actual income is unreliable, we used these household quality measures as a proxy for income [41]. An additional

income measure was the type of ration card held by each individual. Ration cards are used in India to guarantee government subsidies for food depending upon income, which we categorized as BPL (below poverty level), ABL (above poverty level), and not holding a card.

Our outcome variable was a binary measure of household latrine ownership.

Network measures

A name generator is the survey instrument used in social network data collection to elicit the important ties of individuals [42]. In this study respondents (termed here ***egos***) were asked to name up to 8 individuals (termed here ***alters***) for each name generator, and there were 12 name generators administered. These included asking respondents who they: 1) borrow money from, 2) give advice to, 3) help with a decision, 4) borrow kerosene or rice from, 5) lend kerosene or rice to, 6) lend money to, 7) obtain medical advice from, 8) engage socially with, 9) are related to, 10) go to temple with, 11) invite to one's home, 12) visit in another's home.

Because previous research has shown that networks derived from multiple name generators can more successfully measure network characteristics related to network composition [43], we collated data from all 12 name generators to create one comprehensive network. Each network tie indicated that one individual had named the other (or vice versa) in at least one name generator. Dyads in which the individual and her social contact were from the same household were excluded from the analysis. For each individual we then created one continuous measure representing the proportion of her alters that own a latrine.

While many multi-level community studies utilize clusters of randomly sampled individuals from DHS data to represent communities, for this study we used groups that are potentially more socially meaningful: community subgroups which were

objectively generated utilizing sociocentric social network data. In this context we utilize a method that assumes a pre-existing sub-structure of communities within the network, and then goes about the task of discovering them, with no pre-conception as to how many communities will emerge. Ideally this method will detect dense subgroups within the larger network, with relatively sparse between-group ties.

To detect communities within our sample we applied the "fast greedy community" function in the igraph package for R, which optimizes a quantity known as *modularity* within the network[44]. Modularity is a measured property of a network obtained by calculating the proportion of edges that fall within groups minus that same proportion calculated for a random network[45]. Mathematically, modularity is defined as the sum of $A_{ij} - k_i k_j / 2m$ where i and j index all vertices that fall within the same community, A_{ij} are the elements of the full network expressed as an adjacency matrix, $k_i k_j$ are the degrees of the vertices, and m is the total number of connected vertices in the network[29]. The larger the value, the more likely it is that the observed network is not due to chance[29].

One can use the modularity function to discover communities as the result of an iterative process in which, one by one, nodes are grouped with other nodes, with modularity calculated at each iteration. The configuration with the highest modularity score is retained, and the procedure is repeated until all nodes are placed in the same community. The modularity scores at each iteration are then compared, and the one that produced the maximum value is used to identify the total number of communities in the network and to assign each node to its respective community. Modularity scores can be positive, suggesting the presence of community structure, or negative, suggesting that the network is less clustered than would be expected at random. The function, therefore, is free to refuse to create any divisions, if the

algorithmic result suggests that there are none. Furthermore the removal of strongly connected nodes from their appropriate groups will incur a high cost in the modularity score, whereas the removal of weakly connected nodes will not[29]. Finally, the modularity function clearly signals when subdividision is complete: if further division of the network does not increase the modularity score then it is not necessary to continue dividing. In our analysis, the modularity function would optimally terminate with a large positive value, indicating the significant presence of communities within the network.

Community and village level measures

Using the network-generated communities we then calculated community level measures using variables that have been shown to be predictive of latrine ownership at the individual level [36]: proportion of community that are higher caste (OBC or general), mean education level, proportion of community that is above poverty level, proportion of community with RCC roofing, proportion of community with private electricity, mean community age, mean number of rooms per house, mean number of beds per house, and proportion of community that are Hindu. We also included a measure of mean age of adults within the community. All of these measures were calculated using non-self values, meaning that the value for each individual was a measure for that individual's community excluding their own value. Non-self means are used for multi-level analysis to avoid possible confounding which can occur when an individual's own measure is included at both the individual and community levels in the same analysis [20 46]. In this sample, communities that had 10 or less individuals were excluded from the analysis in order to ensure more reliable community level means [20]. All higher-level measures were then replicated at the village level.

Statistical Analyses

We first generated separate models for alter latrine proportion, community latrine proportion, and village latrine proportion using logistic regression including demographic and household covariates.

Next our multilevel analysis included variables at 3 nested levels: village, community, and individual. We estimated 3-level hierarchical logistic regression models to adjust for the clustering of observations at the community and village levels, and we partitioned the variance of the dependent variables into individual, community, and village components. We included predictor variables at each level as fixed effects.

Human subjects

This analysis was conducted with no identifying information for the subjects involved. The institutional review board at the University of California San Diego approved the research protocol for this study.

Results

There were 16984 initial participants in the individual survey. After excluding observations with missing data, and removing community clusters of 10 or fewer people our dataset included 16417 individual participants in 392 communities within 75 villages. The mean community size was 53 (SD 27) with a range of 11 to 185. The mean number of respondents per village was 243 (SD 73) with a range of 90 to 398. Table 4.1 shows descriptive statistics for the study population. Approximately 25% of participants were scheduled caste, and 6% scheduled tribe, both groups that have been marginalized socially and economically. Mean rate of latrine ownership within the sample was 30%, although the range varied according to village (range 0.02 to 0.58) and community (range 0.00 to 0.76).

Figure 4.1 shows the distribution of latrine ownership proportions at the village versus community levels. Although these two measures are correlated (r =.78), note that within villages, communities can vary widely in the proportion of their members who own latrines. Table 4.2 shows the range of values for all measures aggregated at the community and village levels. In addition to the latrine measures, nearly all other measures show greater between-group variation at the community level than at the village level, suggesting that communities may be better units of analysis for discovering reference groups.

To see whether latrine ownership is correlated between neighbors, within communities, and within villages, we perform logistic regression analysis for each of these variables. Table 4.3 shows strongly significant and positive associations between each of our social group latrine proportion variables, controlling for individual and household level demographic covariates. For each standard deviation increase in the proportion of direct social contacts with latrines the odds of individual latrine ownership increase by 1.59 (95% CI 1.56-1.63), compared to 1.84 (95% CI 1.80-1.88) at the community level and 1.88 (95%CI 1.84-1.93) at the village level.

To assess to what degree higher-level random effects accounted for variance at the outcome level, we also calculated the intra-class correlation (ICC) for an unconditional model that included only the higher-level random effects (not shown). These initial estimates suggest that 13% of the variance of individual latrine ownership can be accounted for at the village level, and 18% at the community level.

In Table 4.4, we analyze several multilevel models that include random effects at the community and village level. Model 1 includes individual level effects and confirms previous analyses that people of a higher caste, higher levels of education, and greater levels of financial resources are more likely to own a latrine (coefficients

for individual level controls in SA table SA4.19). In Model 2 we add a variable for the proportion of social contacts who own a latrine. The results of this model show that an increase of 1 standard deviation in the proportion of friends with a latrine increases the odds of individual latrine ownership by 1.27 (95% CI 1.21-1.33), suggesting that latrine ownership is locally clustered in the network. Moreover, this variable remains significant and at about the same magnitude in all models.

For Model 3, we first generated 18 separate models in which, one by one, we added each of the community and village level variables to Model 2 (full models shown in SA tables SA4.1-SA4.18). Of those, the mean age of village respondents, the proportion of the village that are Hindu, the proportion of the community that are Hindu, the community mean number of rooms within households, the mean level of community education, and the mean age of community respondents were significant at a *p* value of 0.10 or less. These significant variables were then added together with the Model 2 variables to create Model 3. Even with the addition of these community and village level variables, the proportion of friends with a latrine remains significant with no change in effect size.

Model 4 includes measures of latrine ownership at the community and village level. While both the village latrine proportion variable and the community latrine proportion variable were strongly significant, the effect size of community latrine proportion was significantly higher. For each one standard deviation increase in proportion of village latrine ownership the odds of individual latrine ownership increase by 1.22 times (95% CI 1.17-1.29) compared with a 1.51 times (95% CI 1.45-1.56) increase in the odds with each standard deviation increase in community proportion (See Figure 4.2). Note also that the community proportion was significantly more related to individual latrine ownership than the social contact

proportion, suggesting that the best reference group for an individual extends beyond their directly-connected social contacts to the broader network in which they are embedded.

As a final robustness check, in Model 5 we remove village level latrine proportion and mean age within the village in order to include village level fixed effects. This is a conservative test of the community level latrine proportion variable because it accounts for any possible factor associated with latrine ownership that might vary at the village level. The model shows that the proportion of latrine owners in the network-defined community retains significance even after controlling for all possible village level influences. A 1 standard deviation increase in community latrine ownership proportion increases the odds of individual latrine ownership by 1.34 (95% CI 1.24-1.44).

Discussion

This study analyzed data from rural India to test social network predictors of individual latrine ownership at 3 levels: that of the individual's direct social contacts, her network community, and the village. Unlike many papers that derive community units from DHS sampling clusters, we used a social network method to determine each participant's community membership from respondent reported networks bounded at the village level. This provided us with a uniquely meaningful measure of community, and one in which individuals could be precisely nested within one unique community within one specific village. Using multilevel modeling, we tested the association between latrine ownership and a series of individual level characteristics as well as community and village level characteristics that were calculated by aggregating measures from the individual level. This is one of the few studies of

which we are aware that have used network-generated communities with multilevel modeling techniques to assess community effects on individual outcomes.

Our analysis shows that for a variety of variables – including latrine adoption – there is greater between-group variation at the community level then there is at the village level. For instance while there are 14 communities in which every single community member has 0 years of education, and some other communities with a disproportionate number of highly educated people. This suggests that network-derived communities are more demographically homogenous than are geographically-defined villages, and may therefore be more representative of the sort of reference group within which norms are held. Furthermore our initial analysis of the random effects model showed us that the community level accounted for a greater proportion of the variance in latrine ownership than did the village level.

When we modeled individual latrine ownership, we found a significant relationship with latrine ownership by others at three different levels. The likelihood of ownership increases with 1) the proportion of latrines owned by one's direct social contacts, 2) the proportion owned by individuals within one's network-defined community, and 3) the proportion owned by individuals within one's village. As predicted, village level latrine ownership showed a higher correlation than social contact latrine ownership. Contrary to our prediction however, we found that community latrine ownership not only predicts individual latrine ownership, but that it does so significantly better than either direct social contact latrine ownership or village level latrine ownership (see Figure 2).

To explain this, we must consider the possibility of *homophily*[47 48], or the fact that individuals may choose their social relations based upon similarities that would also lead each individual to build a latrine. While people are likely to choose their

direct social contacts based upon similarities that could be associated with latrine ownership, such as being of the same caste or same level of education, in our analysis this level of effect was the least prominent. It is possible that people could choose villages based upon shared characteristics that would also predict latrine ownership, however in the context of rural India such between-village mobility is not common[49]. Typically families stay within the same villages for generations. And although communities could also be formed from commonalities that predict latrine ownership, the most significant of these were controlled for at both the individual, community, and village level, including a model with fixed effects that controls for all possible village-level effects. The fact that there is significance at all three levels of the model suggests a social dynamic that is consistent, strong, and potentially complex.

We had expected that the village level effect in this analysis would be the strongest based on the fact that latrine-building interventions are typically implemented at the village level. What does it tell us that socially determined communities, derived from an objective mathematical algorithm, are more predictive of latrine ownership than either direct social contacts, or those geographically proximal within the village? We conjecture that these network-determined communities actually make up the salient reference groups to which individuals turn for information on normative expectations. Descriptive norms can potentially diffuse between people within the same geographic vicinity, such as between strangers at the local market[50]. Injunctive norms, however, can only take place within communities that are socially relevant to an individual, and with whom that individual strives to maintain positive relations[16]. The possible social effects we observed in this sample could certainly be the result of both dynamics. However, the fact that socially

connected network communities are the most significantly correlated with latrine ownership suggests that an injunctive normative dynamic may be the most significant for latrine adoption.

Finally, we should emphasize the utility of using network communities (rather than villages) to inform health intervention efforts. Since traditional latrine-building interventions target village units, we might expect that village level latrine adoption would be the best predictor of individual level latrine adoption, and community level effects, if any, would be subordinate. However our results show the opposite. If the village unit is not the primary source of normative expectations for individuals, village level interventions aimed at changing norms may be inefficient. Future work should explore interventions that focus on identifying and engaging socially meaningful community groups within the village in order to see whether this helps maximize the likelihood of widespread and enduring behavior change.

Limitations of this study

There are several limitations to this study. First, because the data is static, we are not able to track time dependent effects that might shed light on causal explanations for the associations we observe. While we conclude that our results are suggestive of causal effects, longitudinal studies with randomized treatment groups would substantially strengthen this analysis. Second, our dataset included the network of approximately 50% of randomly sampled adults in each village, so the missing half of the network, although randomly omitted, could potentially change the impact of our findings. Finally, latrine ownership does not guarantee latrine usage. Further analysis could shed light on the degree to which latrine usage is socially predicted, and whether network based interventions eventually lead to a decrease in morbidity and mortality from sanitation related infectious disease.

Conclusion

Work on higher-level community effects is proliferating among health behavior researchers as we advance our understanding of the reciprocal relationship between the individual and his/her greater social environment. An individually focused perspective is particularly deficient when attempting to understand normatively driven behaviors, which are heavily dependent upon the expectations of salient others. While qualitative work has shed light on many of these dynamics, quantitative analysis has the ability to provide a more structured perspective on how social effects contribute to individual behavior. This research contributes to that ongoing process, by illuminating a potentially powerful social determinant of latrine adopting behavior, the social relevant community to which an individual belongs.

Table 4.1. Demographic and individual characteristics of adult respondents from rural Karnataka India N=16417	
	Mean (SD)
Age in years	39.30 (12.5)
Education mean years	5.02 (4.65)
Number of rooms in house	2.62 (1.54)
Number of beds in house	1.05 (1.62)
Social contact latrine proportion	0.30 (0.19)
	Proportion of total sample
Gender (% female)	0.551
Caste	
Scheduled caste	0.252
Scheduled tribe	0.060
OBC	0.566
General	0.122
Ration card category	
Above poverty	0.175
Below poverty	0.669
Does not own card	0.157
Roof-type	
Thatch	0.020
Tile	0.314
Stone	0.305
Sheet	0.182
RCC	0.138
Other	0.041
Household electricity	
Private	0.657
Govt.	0.285
No	0.058
Latrine ownership yes	0.304

Table 4.2 Community and village level attributes aggregated from individual measures of respondents from rural Karnataka India N=16417

	Overall score (Mean or proportion across entire sample)	Community level Between Group Variation SD(Range)	Village level Between Group Variation SD(Range)
Proportion with household latrine	0.30	0.16 (0.00-.77)	0.12 (0.02-0.58)
Proportion high caste	0.69	0.22 (0.00-1.00)	0.16 (0.04-1.00)
Proportion above poverty line	0.17	0.12 (0.00-0.57)	0.09 (0.00-0.39)
Proportion Hindu	0.96	0.10 (0.25-1.00)	0.08 (0.59-1.00)
Proportion with RCC roofing	0.14	0.11 (0.00-0.56)	0.09 (0.00-0.45)
Proportion with household elec	0.66	0.18 (0.00-1.00)	0.14 (0.9-0.89)
Mean age in years	39.1	2.44 (23.00-47.00)	1.35 (35.90-42.30)
Mean education in years	5.02	1.34 (0.75-10.00)	0.97(2.20-6.69)
Mean number of beds in house	1.05	0.72 (0.00-6.67)	0.59 (0.27-3.25)
Mean number of rooms in house	2.62	0.54 (1.00-6.17)	0.38 (1.56-3.46)

**Means and proportions using non-self measures*

Table 4.3. Results of logistic regression analysis showing effect of social group latrine ownership proportions on individual latrine ownership at the levels of the individual, the community, and the village									
	Model 1: level 1 Social contact latrine proportion			Model 2: level 2 Community latrine proportion			Model 3: level 3 Village and community characteristics		
Fixed Part	Beta	SE	P	Beta	SE	P	Beta	SE	P
Intercept	-2.69	0.25		-3.45	0.26		4.07	2.52	
Proportion of contacts with a latrine	**1.61**	**0.07**	**<.001**						
Community latrine proportion				**3.82**	**0.14**	**<.001**			
Village latrine proportion							**5.09**	**0.19**	**<.001**
Caste									
Scheduled caste	Ref			Ref			Ref		
Scheduled tribe	-0.32	0.12	0.01	-0.25	0.12	0.04	-0.12	0.12	0.30
OBC	0.44	0.06	<.001	0.64	0.06	<.001	0.74	0.06	<.001
General	0.39	0.08	<.001	0.69	0.08	<.001	0.84	0.08	<.001
HH electricity									
Private	Ref			Ref			Ref		
Govt.	-0.93	0.06	<.001	-0.90	0.06	<.001	-0.89	0.06	<.001
No	-1.45	0.16	<.001	-1.48	0.16	<.001	-1.47	0.16	<.001
# of rooms in house	0.29	0.02	<.001	0.30	0.02	<.001	0.31	0.02	<.001
# of beds in house	0.20	0.02	<.001	0.17	0.02	<.001	0.17	0.02	<.001
Roof-type									
Thatch	Ref			Ref			Ref		
Tile	-0.15	0.24	0.52	-0.22	0.24	0.37	-0.21	0.24	0.38
Stone	0.45	0.24	0.06	0.42	0.24	0.08	0.41	0.24	0.09
Sheet	0.45	0.24	0.06	0.38	0.24	0.12	0.32	0.24	0.19
RCC	1.27	0.24	<.001	1.27	0.25	<.001	1.26	0.25	<.001
Other	-0.23	0.26	0.39	-0.30	0.27	0.27	-0.32	0.27	0.23
Years of education	-0.05	0.00	<.001	0.05	0.00	<.001	0.06	0.00	<.001
Ration card									
Above poverty line	Ref			Ref			Ref		
Below poverty line	-0.63	0.05	<.001	-0.71	0.05	<.001	-0.72	0.05	<.001
No ration card	-0.33	0.07	<.001	-0.38	0.07	<.001	-0.37	0.07	<.001

Table 4.4 Results of multi-level logistic regression analysis showing the effect of social group latrine ownership proportions on individual latrine ownership at the levels of the individual, the community, and the village*

	Model 1 Individual attributes	Model 2 Model 1 plus social contact latrine	Model 3 Model 2 plus village and community characteristics	Model 4 Model 3 plus community and village latrine ownership	Model 5 community latrine plus village level fixed effects (not shown)
Fixed Part	Beta (SE)	Beta (SE)	Beta (SE)	Beta (SE)	Beta (SE)
Intercept	-2.75 (0.29)	-2.9 (0.28)	-2.885 (0.28)	-2.773 (0.264)	-2.337*** (0.324)
Proportion of contacts with a latrine		**0.81*** (0.08)**	**0.806*** (0.082)**	**0.823 (0.081)**	**0.743*** (0.081)**
Community latrine proportion				**2.58 (0.245)**	**1.782*** (0.258)**
Village latrine proportion				**1.683 (0.422)**	
Village level variance	0.72	0.54	0.48	.08	
Community level variance	0.20	0.17	0.16	>.001	>.001
AIC	13893	13806	13799	13769	13695
LL	-6929	-6884	-6874	-6858	-6750

*Individual, community and village level controls now shown. Full models shown in SA table SA4.19

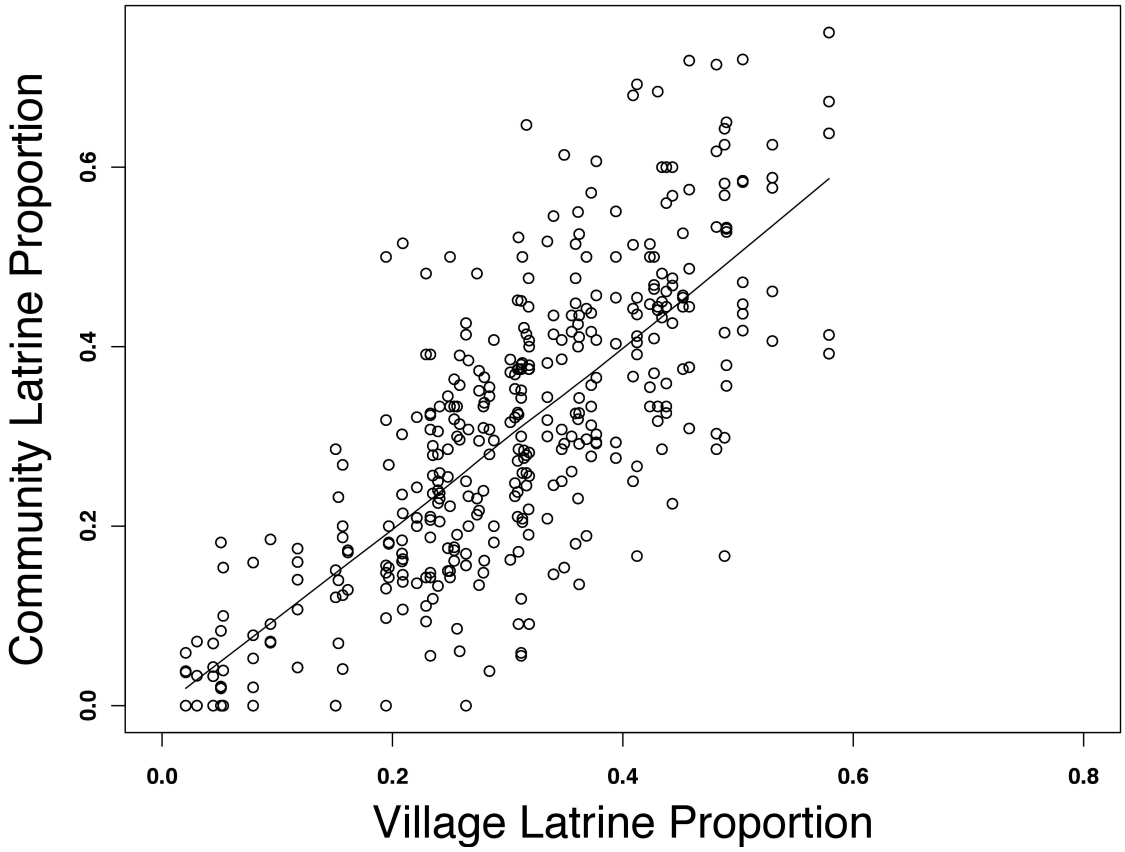

Figure 4.1 The village latrine proportion by community latrine proportion. Although overall there is a positive relationship between the proportion of people in a community with latrines and the proportion of people in that community's village with latrines, we can see from this figure that there is considerable variation as well.

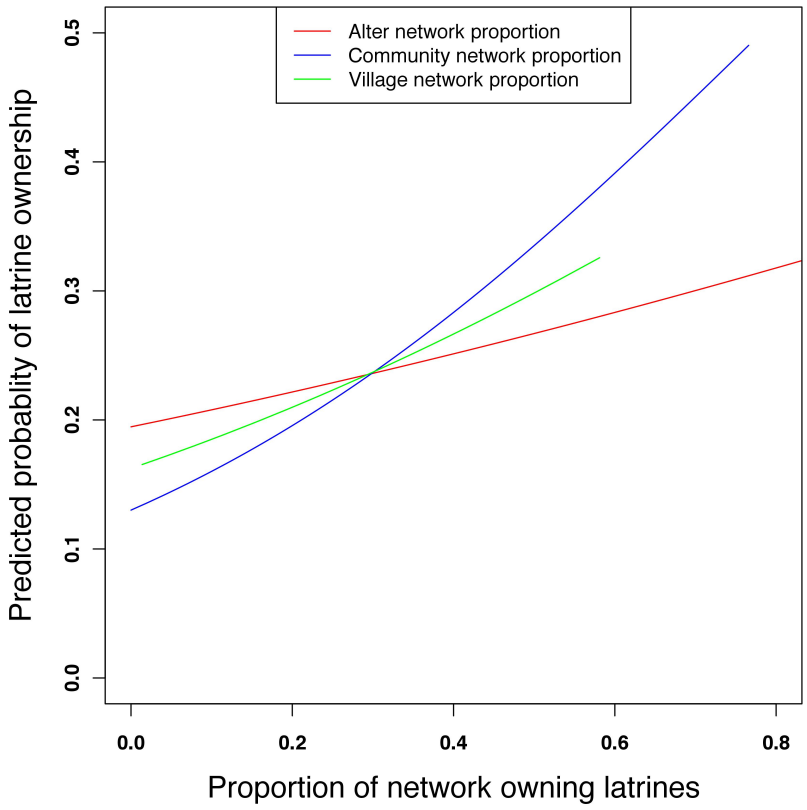

Figure 4.2. The predicted probability of latrine ownership by the proportion of the village network, (green), community network (blue) and the network of direct social contacts, termed here alters, (red) that own latrines, holding all other variables in the model constant (mean values or most prevalent category). We can see here that the community latrine ownership proportion is has the strongest correlation with individual latrine ownership.

References

1. WHO, UNICEF. Meeting the MGD drinking water and sanitation target: the urban and rural challenge of the decade. In: WHO, UNICEF, editors. Switzerland, 2006.

2. Pruss A, Kay D, Fewtrell L, Bartram J. Estimating the burden of disease from water, sanitation, and hygiene at a global level. *Environmental health perspectives* 2002;110(5):537-42.

3. Pattanayak SK, Yang JC, Dickinson KL, Poulos C, Patil SR, Mallick RK, et al. Shame or subsidy revisited: social mobilization for sanitation in Orissa, India. *Bulletin of the World Health Organization* 2009;87(8):580-87

4. Black RE, Cousens S, Johnson HL, Lawn JE, Rudan I, Bassani DG, et al. Global, regional, and national causes of child mortality in 2008: a systematic analysis. *The Lancet* 2010;375(9730):1969-87.

5. Dickinson K, Pattanayak S. Open Sky Latrines: Do social effects influence technology adoption in the case of a (very) impure public good? . *Working Paper, Duke University* 2012

6. U.N. The Millennium Development Goals Report. In: Nations U, editor. New York New York, 2011

7. Bonu S, Kim H. Sanitation in India: Progress, Differentials, Correlates, and Challenges. In: Bank AD, editor. *South Asia Occasional Paper Series No. 2*. Mandaluyong City: Asian Development Bank, 2009.

8. Dyalchand A, Khale M, Vasudevan S, Kale N. What Communication and Institutional Arrangements Influence Sanitation Related Social Norms in Rural India? In: L M, S M, editors. *Shit Matters: The Potential of Community-Led Total Sanitation* London: Practical Action, 2008.

9. Dubois N, editor. *A Sociocognitive Approach to Social Norms* Bristol Routledge, 2003

10. Bicchieri C. *The Grammar of Society: the Nature and Dynamics of Social Norms*. New York: Cambridge University Press, 2006.

11. Cialdini RB, Reno RR, Kallgren CA. A focus theory of normative conduct: Recycling the concept of norms to reduce littering in public places. *Journal of Personality and Social Psychology* 1990;58(6):1015.

12. Cialdini RB, Kallgren CA, Reno RR. A focus theory of normative conduct: A theoretical refinement and reevaluation of the role of norms in human behavior. *Advances in experimental social psychology* 1991;24(20):1-243.

13. Coleman JS. The emergence of norms. In: Wippler R, Opp K-D, Hecther M, editors. *Social Institutions: Their Emergence, Maintenance, and Effects*. New York: Aldine de Gruyter, 1990:35-39.

14. Bicchieri C. The rules we live by. In: Guala Francesco; Steel D, editor. *The Philosophy of Social Science Reader*: Routledge, 2010.

15. Bendor J, Swistak P. The evolution of norms. *American Journal of Sociology* 2001:1493-545.

16. Jacobson RP, Mortensen CR, Cialdini RB. Bodies obliged and unbound: Differentiated response tendencies for injunctive and descriptive social norms. *Journal of Personality and Social Psychology* 2011;100(3):433.

17. Cialdini RB. Descriptive social norms as underappreciated sources of social control. *Psychometrika* 2007;72(2):263-68.

18. LeJeune J, Mackie G. Social dynamics of abandonment of harmful practices: a new look at the theory. In: UNICEF, editor. Florence: UNICEF Innocenti Research Centre, 2008.

19. Bicchieri C, Muldoon R. Social Norms. In: Zalta EN, editor. *Stanford Encyclopedia of Philosophy*. Spring 2011 ed. Stanford CA: The Metaphysics Research Lab 2011.

20. Storey J, Kaggwa E. The Influence of Changes in Fertility Related Norms on Contraceptive Use in Egypt, 1995-2005. *Population Review* 2009;48(1):1-19.

21. Kaggwa E, Diop N, Storey J. The role of individual and community normative factors: A multilevel analysis of contraceptive use among women in Mali. *International family planning perspectives* 2008;34(2):79-88.

22. Hayford SR. Conformity and change: community effects on female genital cutting in Kenya. *Journal of Health and Social Behavior* 2005;46(2):121-40 %@ 0022-1465.

23. Boyle MH, Georgiades K, Cullen J, Racine Y. Community influences on intimate partner violence in India: Women's education, attitudes towards mistreatment and standards of living. *Social Science & Medicine* 2009;69(5):691-97 %@ 0277-9536.

24. Bernburg JG, Thorlindsson T. Violent values, conduct norms, and youth aggression: A multilevel study in Iceland. *Sociological quarterly* 2005;46(3):457-78 %@ 1533-8525.

25. Barrientos-Gutierrez T, Gimeno D, Mangione TW, Harrist RB, Amick BC. Drinking social norms and drinking behaviours: a multilevel analysis of 137 workgroups in

16 worksites. *Occupational and environmental medicine* 2007;64(9):602-08 %@ 1470-7926.

26. Wiium N, Torsheim T, Wold B. Normative processes and adolescents' smoking behaviour in Norway: A multilevel analysis. *Social Science & Medicine* 2006;62(7):1810-18 %@ 0277-9536.

27. Porter MA, Onnela JP, Mucha PJ. Communities in networks. *Notices of the AMS* 2009;56(9):1082-97.

28. Moody J. Peer influence groups: identifying dense clusters in large networks. *Social Networks* 2001;23(4):261-83 %@ 0378-8733.

29. Newman MEJ. Modularity and community structure in networks. *Proceedings of the National Academy of Sciences* 2006;103(23):8577-82.

30. Kreager DA, Rulison K, Moody J. DELINQUENCY AND THE STRUCTURE OF ADOLESCENT PEER GROUPS*. *Criminology* 2011;49(1):95-127 %@ 1745-9125.

31. Cleveland MJ, Feinberg ME, Osgood DW, Moody J. Do Peers' Parents Matter? A New Link Between Positive Parenting and Adolescent Substance Use. *Journal of studies on alcohol and drugs* 2012;73(3):423.

32. Cruz JE, Emery RE, Turkheimer E. Peer Network Drinking Predicts Increased Alcohol Use From Adolescence to Early Adulthood After Controlling for Genetic and Shared Environmental Selection. 2012.

33. Grzywacz JG, Fuqua J. The social ecology of health: leverage points and linkages. *Behavioral Medicine* 2000;26(3):101-15.

34. Stokols D. Social ecology and behavioral medicine: implications for training, practice, and policy. *Behavioral Medicine* 2000;26(3):129-38.

35. Jenkins MW, Curtis V. Achieving the [] good life': Why some people want latrines in rural Benin. *Social Science & Medicine* 2005;61(11):2446-59 %@ 0277-9536.

36. Shakya H, Christakis NA, Fowler JH. Social network predictors of latrine ownership in rural India. 2012.

37. Chambers R. Going to Scale with Community-Led Total Sanitation: Reflections on Experience, Issues and Ways Forward. *IDS Practice Papers* 2009;2009(1):01-50 %@ 2040-0225.

38. Banerjee A, Chandrasekhar AG, Duflo E, Jackson MO. The Diffusion of Microfinance: Mimeo. http://econ-www. mit. edu/files/6910, 2011.

39. Jackson MO, Rodriguez-Barraquer T, Tan X. Social patterns and social quilts: network patterns of favor exchange. *American Economic Review* 2012.

40. Marin A, Wellman B. Social network analysis: An introduction. *The SAGE Handbook of Social Network Analysis* 2011:11.

41. Morris SS, Carletto C, Hoddinott J, Christiaensen LJM. Validity of rapid estimates of household wealth and income for health surveys in rural Africa. *Journal of epidemiology and community health* 2000;54(5):381-87 %@ 1470-2738

42. Marin A. Are respondents more likely to list alters with certain characteristics?Implications for name generator data. *Social Networks* 2004.

43. Marin A, Hampton KN. Simplifying the Personal Network Name Generator. *Field Methods* 2007;19(2):163.

44. igraph: Network analysis and visualization [program]. 0.5.4-1 version, 2010.

45. Clauset A, Newman MEJ, Moore C. Finding community structure in very large networks. *Physical review E* 2004;70(6):066111.

46. Storey J, Schoemaker J. Communication, normative influence and the sustainability of health behavior over time: A multilevel analysis of contraceptive use in Indonesia, 1997-2003 *Health Communication Division of the International Communication Association, 55th Annual Conference*. Dresden Germany, 2006.

47. McPherson M, Smith-Lovin L, Cook JM. Birds of a feather: Homophily in social networks. *Annual Review of Sociology* 2001:415-44 %@ 0360-572.

48. Aral S, Muchnik L, Sundararajan A. Distinguishing influence-based contagion from homophily-driven diffusion in dynamic networks. *Proceedings of the National Academy of Sciences* 2009;106(51):21544-49.

49. Munshi K, MR R. Why Is Mobility in India so Low? Social Insurance, Inequality, and Growth. *CID Working Paper #121* 2005.

50. Kohler HP, Behrman JR, Watkins SC. The density of social networks and fertility decisions: Evidence from South Nyanza District, Kenya. *Demography* 2001;38(1):43-58 %@ 0070-3370.

Supplementary Appendix

Table SA 4.1: Results of logistic regression analysis showing effect of proportion of community high caste on individual latrine ownership

Fixed Part	Beta	SE	P
Intercept	-2.895	0.282	0.000
Proportion of community high caste	-0.170	0.204	0.406
Proportion of contacts with a latrine	0.808	0.082	< 2e-16
Caste			
Scheduled caste	Ref		
Scheduled tribe	0.035	0.131	0.790
OBC	0.673	0.071	< 2e-16
General	0.862	0.096	< 2e-16
HH electricity			
Private	Ref		
Govt.	-0.93	0.06	<.001
No	-1.45	0.16	<.001
# of beds in house	0.190	0.018	< 2e-16
# of rooms in house	0.349	0.019	< 2e-16
Roof-type			
Thatch	Ref		
Tile	-0.319	0.256	0.213
Stone	0.529	0.258	0.040
Sheet	0.466	0.258	0.070
RCC	1.437	0.261	0.000
Other	-0.245	0.284	0.389
Years of education	0.208	0.018	< 2e-16
Ration card			
Above poverty line	Ref		
Below poverty line	-0.713	0.059	< 2e-16
No ration card	-0.379	0.073	0.000
Village level variance	0.54		
Community level variance	0.17		
AIC	13808		
LL	-6884		

Table SA 4.2: Results of logistic regression analysis showing effect of proportion of community above poverty line on individual latrine ownership

Fixed Part	Beta	SE	P
Intercept	-2.886	0.282	< 2e-16
Proportion of community above poverty line	0.057	0.387	0.882
Proportion of contacts with a latrine	0.807	0.082	< 2e-16
Caste			
Scheduled caste	Ref		
Scheduled tribe	0.026	0.131	0.840
OBC	0.659	0.069	< 2e-16
General	0.848	0.095	< 2e-16
HH electricity			
Private	Ref		
Govt.	-0.957	0.063	< 2e-16
No	-1.467	0.165	< 2e-16
# of beds in house	0.190	0.018	< 2e-16
# of rooms in house	0.349	0.019	< 2e-16
Roof-type			
Thatch	Ref		
Tile	-0.318	0.256	0.214
Stone	0.532	0.258	0.039
Sheet	0.468	0.258	0.070
RCC	1.439	0.261	0.000
Other	-0.243	0.284	0.394
Years of education	0.208	0.018	< 2e-16
Ration card			
Above poverty line	Ref		
Below poverty line	-0.712	0.060	< 2e-16
No ration card	-0.378	0.073	0.000
Village level variance	0.54		
Community level variance	0.17		
AIC	13808		
LL	-6884		

Table SA 4.3: Results of logistic regression analysis showing effect of mean community education level on individual latrine ownership

Fixed Part	Beta	SE	P
Intercept	-2.878	0.281	< 2e-16
Mean community education level	0.064	0.033	0.055
Proportion of contacts with a latrine	0.800	0.082	< 2e-16
Caste			
Scheduled caste	Ref		
Scheduled tribe	0.021	0.131	0.874
OBC	0.654	0.069	< 2e-16
General	0.838	0.095	< 2e-16
HH electricity			
Private	Ref		
Govt.	-0.955	0.063	< 2e-16
No	-1.467	0.165	< 2e-16
# of beds in house	0.189	0.018	< 2e-16
# of rooms in house	0.348	0.019	< 2e-16
Roof-type			
Thatch	Ref		
Tile	-0.321	0.256	0.209
Stone	0.530	0.258	0.040
Sheet	0.464	0.258	0.071
RCC	1.435	0.261	0.000
Other	-0.244	0.284	0.391
Years of education	0.209	0.018	< 2e-16
Ration card			
Above poverty line	Ref		
Below poverty line	-0.713	0.059	< 2e-16
No ration card	-0.377	0.073	0.000
Village level variance	0.53		
Community level variance	0.17		
AIC	13805		
LL	-6882		

Table SA 4.4: Results of logistic regression analysis showing effect of proportion of community with RCC roof on individual latrine ownership

Fixed Part	Beta	SE	P
Intercept	-2.889	0.282	< 2e-16
Proportion of community with RCC roof	-0.501	0.385	0.193
Proportion of contacts with a latrine	0.811	0.082	< 2e-16
Caste			
Scheduled caste	Ref		
Scheduled tribe	0.030	0.131	0.819
OBC	0.661	0.069	< 2e-16
General	0.849	0.095	< 2e-16
HH electricity			
Private	Ref		
Govt.	-0.958	0.063	< 2e-16
No	-1.468	0.165	< 2e-16
# of beds in house	0.190	0.018	< 2e-16
# of rooms in house	0.349	0.019	< 2e-16
Roof-type			
Thatch	Ref		
Tile	-0.320	0.256	0.210
Stone	0.532	0.258	0.039
Sheet	0.468	0.258	0.070
RCC	1.440	0.261	0.000
Other	-0.242	0.284	0.394
Years of education	0.208	0.018	< 2e-16
Ration card			
Above poverty line	Ref		
Below poverty line	-0.713	0.059	< 2e-16
No ration card	-0.378	0.073	0.000
Village level variance	0.54		
Community level variance	0.17		
AIC	13807		
LL	-6883		

Table SA 4.5: Results of logistic regression analysis showing effect of proportion of community with private electricity on individual latrine ownership

Fixed Part	Beta	SE	P
Intercept	-2.889	0.282	< 2e-16
Proportion of community with private electricity	-0.211	0.282	0.453
Proportion of contacts with a latrine	0.809	0.082	< 2e-16
Caste			
Scheduled caste	Ref		
Scheduled tribe	0.032	0.131	0.809
OBC	0.663	0.069	< 2e-16
General	0.854	0.095	< 2e-16
HH electricity			
Private	Ref		
Govt.	-0.963	0.064	< 2e-16
No	-1.472	0.165	< 2e-16
# of beds in house	0.190	0.018	< 2e-16
# of rooms in house	0.349	0.019	< 2e-16
Roof-type			
Thatch	Ref		
Tile	-0.319	0.256	0.213
Stone	0.530	0.258	0.040
Sheet	0.467	0.258	0.070
RCC	1.438	0.261	0.000
Other	-0.244	0.284	0.391
Years of education	0.208	0.018	< 2e-16
Ration card			
Above poverty line	Ref		
Below poverty line	-0.712	0.059	< 2e-16
No ration card	-0.379	0.073	0.000
Village level variance	0.54		
Community level variance	0.16		
AIC	13807		
LL	-6884		

Table SA 4.6: Results of logistic regression analysis showing effect of community mean number of beds within a household on individual latrine ownership			
Fixed Part	Beta	SE	P
Intercept	-2.886	0.282	< 2e-16
Community mean number of beds in household	0.002	0.076	0.981
Proportion of contacts with a latrine	0.807	0.082	< 2e-16
Caste			
Scheduled caste	Ref		
Scheduled tribe	0.027	0.131	0.838
OBC	0.659	0.069	< 2e-16
General	0.848	0.095	< 2e-16
HH electricity			
Private	Ref		
Govt.	-0.957	0.063	< 2e-16
No	-1.466	0.165	< 2e-16
# of beds in house	0.190	0.018	< 2e-16
# of rooms in house	0.349	0.019	< 2e-16
Roof-type			
Thatch	Ref		
Tile	-0.318	0.256	0.214
Stone	0.532	0.258	0.039
Sheet	0.468	0.258	0.070
RCC	1.439	0.261	0.000
Other	-0.243	0.284	0.394
Years of education	0.208	0.018	< 2e-16
Ration card			
Above poverty line	Ref		
Below poverty line	-0.713	0.059	< 2e-16
No ration card	-0.379	0.073	0.000
Village level variance	0.54		
Community level variance	0.17		
AIC	13808		
LL	-6884		

Table SA 4.7: Results of logistic regression analysis showing effect of mean number of rooms within a household on individual latrine ownership

Fixed Part	Beta	SE	P
Intercept	-2.903	0.282	< 2e-16
Mean community number of rooms in household	-0.150	0.086	0.082
Proportion of contacts with a latrine	0.815	0.082	< 2e-16
Caste			
Scheduled caste	Ref		
Scheduled tribe	0.028	0.131	0.830
OBC	0.663	0.069	< 2e-16
General	0.851	0.095	< 2e-16
HH electricity			
Private	Ref		
Govt.	-0.957	0.063	< 2e-16
No	-1.467	0.165	< 2e-16
# of beds in house	0.191	0.018	< 2e-16
# of rooms in house	0.353	0.019	< 2e-16
Roof-type			
Thatch	Ref		
Tile	-0.318	0.256	0.213
Stone	0.529	0.258	0.040
Sheet	0.466	0.258	0.071
RCC	1.436	0.261	0.000
Other	-0.246	0.284	0.388
Years of education	0.208	0.018	< 2e-16
Ration card			
Above poverty line	Ref		
Below poverty line	-0.713	0.059	< 2e-16
No ration card	-0.379	0.073	0.000
Village level variance	0.53		
Community level variance	0.17		
AIC	13805		
LL	-6883		

Table SA 4.8: Results of logistic regression analysis showing effect of proportion of the community that is Hindu on individual latrine ownership

Fixed Part	Beta	SE	P
Intercept	-2.887	0.281	< 2e-16
Proportion of community that is Hindu	-1.059	0.536	0.048
Proportion of contacts with a latrine	0.808	0.082	< 2e-16
Caste			
Scheduled caste	Ref		
Scheduled tribe	0.024	0.131	0.854
OBC	0.657	0.069	< 2e-16
General	0.847	0.095	< 2e-16
HH electricity			
Private	Ref		
Govt.	-0.958	0.063	< 2e-16
No	-1.465	0.164	< 2e-16
# of beds in house	0.190	0.018	< 2e-16
# of rooms in house	0.348	0.019	< 2e-16
Roof-type			
Thatch	Ref		
Tile	-0.313	0.256	0.222
Stone	0.536	0.258	0.038
Sheet	0.472	0.258	0.067
RCC	1.441	0.261	0.000
Other	-0.237	0.284	0.405
Years of education	0.208	0.018	< 2e-16
Ration card			
Above poverty line	Ref		
Below poverty line	-0.713	0.059	< 2e-16
No ration card	-0.380	0.073	0.000
Village level variance	0.52		
Community level variance	0.17		
AIC	13804		
LL	-6882		

Table SA 4.9: Results of logistic regression analysis showing effect of mean age of respondents in community on individual latrine ownership

Fixed Part	Beta	SE	P
Intercept	-2.890	0.281	< 2e-16
Mean age of community respondents	-0.034	0.016	0.035
Proportion of contacts with a latrine	0.808	0.082	< 2e-16
Caste			
Scheduled caste	Ref		
Scheduled tribe	0.025	0.131	0.850
OBC	0.659	0.069	< 2e-16
General	0.850	0.095	< 2e-16
HH electricity			
Private	Ref		
Govt.	-0.957	0.063	< 2e-16
No	-1.467	0.165	< 2e-16
# of beds in house	0.190	0.018	< 2e-16
# of rooms in house	0.348	0.019	< 2e-16
Roof-type			
Thatch	Ref		
Tile	-0.316	0.256	0.216
Stone	0.533	0.258	0.039
Sheet	0.467	0.258	0.070
RCC	1.439	0.261	0.000
Other	-0.242	0.284	0.395
Years of education	0.210	0.018	< 2e-16
Ration card			
Above poverty line	Ref		
Below poverty line	-0.711	0.059	< 2e-16
No ration card	-0.375	0.073	0.000
Village level variance	0.53		
Community level variance	0.18		
AIC	13804		
LL	-6882		

Table SA 4.10: Results of logistic regression analysis showing effect of proportion of village that is high caste on individual latrine ownership

Fixed Part	Beta	SE	P
Intercept	-2.886	0.282	< 2e-16
Proportion of village that is high caste	-0.493	0.585	0.399
Proportion of contacts with a latrine	0.807	0.082	< 2e-16
Caste			
Scheduled caste	Ref		
Scheduled tribe	0.029	0.131	0.823
OBC	0.662	0.069	< 2e-16
General	0.851	0.095	< 2e-16
HH electricity			
Private	Ref		
Govt.	-0.958	0.063	< 2e-16
No	-1.466	0.165	< 2e-16
# of beds in house	0.190	0.018	< 2e-16
# of rooms in house	0.349	0.019	< 2e-16
Roof-type			
Thatch	Ref		
Tile	-0.319	0.256	0.213
Stone	0.528	0.258	0.041
Sheet	0.466	0.258	0.071
RCC	1.436	0.261	0.000
Other	-0.245	0.284	0.390
Years of education	0.208	0.018	< 2e-16
Ration card			
Above poverty line	Ref		
Below poverty line	-0.713	0.059	< 2e-16
No ration card	-0.380	0.073	0.000
Village level variance	0.53		
Community level variance	0.17		
AIC	13808		
LL	-6884		

Table SA 4.11: Results of logistic regression analysis showing effect of proportion of village that is above poverty level on individual latrine ownership

Fixed Part	Beta	SE	P
Intercept	-2.887	0.282	< 2e-16
Proportion of village that is above poverty level	-0.198	1.021	0.846
Proportion of contacts with a latrine	0.807	0.082	< 2e-16
Caste			
Scheduled caste	Ref		
Scheduled tribe	0.027	0.131	0.836
OBC	0.659	0.069	< 2e-16
General	0.849	0.095	< 2e-16
HH electricity			
Private	Ref		
Govt.	-0.957	0.063	< 2e-16
No	-1.466	0.165	< 2e-16
# of beds in house	0.190	0.018	< 2e-16
# of rooms in house	0.349	0.019	< 2e-16
Roof-type			
Thatch	Ref		
Tile	-0.318	0.256	0.213
Stone	0.531	0.258	0.040
Sheet	0.467	0.258	0.070
RCC	1.438	0.261	0.000
Other	-0.243	0.284	0.393
Years of education	0.208	0.018	< 2e-16
Ration card			
Above poverty line	Ref		
Below poverty line	-0.713	0.059	< 2e-16
No ration card	-0.379	0.073	0.000
Village level variance	0.53		
Community level variance	0.17		
AIC	13808		
LL	-6884		

Table SA 4.12: Results of logistic regression analysis showing effect of mean village education level on individual latrine ownership

Fixed Part	Beta	SE	P
Intercept	-2.886	0.281	< 2e-16
Mean village education level	0.091	0.095	0.340
Proportion of contacts with a latrine	0.805	0.082	< 2e-16
Caste			
Scheduled caste	Ref		
Scheduled tribe	0.025	0.131	0.850
OBC	0.659	0.069	< 2e-16
General	0.846	0.095	< 2e-16
HH electricity			
Private	Ref		
Govt.	-0.955	0.063	< 2e-16
No	-1.467	0.165	< 2e-16
# of beds in house	0.190	0.018	< 2e-16
# of rooms in house	0.349	0.019	< 2e-16
Roof-type			
Thatch	Ref		
Tile	-0.318	0.256	0.214
Stone	0.532	0.258	0.039
Sheet	0.466	0.258	0.071
RCC	1.438	0.261	0.000
Other	-0.244	0.284	0.390
Years of education	0.209	0.018	< 2e-16
Ration card			
Above poverty line	Ref		
Below poverty line	-0.713	0.059	< 2e-16
No ration card	-0.378	0.073	0.000
Village level variance	0.53		
Community level variance	0.17		
AIC	13807		
LL	-6884		

Table SA 4.13: Results of logistic regression analysis showing effect of proportion of village with RCC roofing			
Fixed Part	Beta	SE	P
Intercept	-2.886	0.282	< 2e-16
Proportion of village with RCC roofing	-0.061	1.069	0.955
Proportion of contacts with a latrine	0.807	0.082	< 2e-16
Caste			
Scheduled caste	Ref		
Scheduled tribe	0.027	0.131	0.837
OBC	0.659	0.069	< 2e-16
General	0.848	0.095	< 2e-16
HH electricity			
Private	Ref		
Govt.	-0.957	0.063	< 2e-16
No	-1.466	0.165	< 2e-16
# of beds in house	0.190	0.018	< 2e-16
# of rooms in house	0.349	0.019	< 2e-16
Roof-type			
Thatch	Ref		
Tile	-0.318	0.256	0.214
Stone	0.532	0.258	0.039
Sheet	0.468	0.258	0.070
RCC	1.439	0.261	0.000
Other	-0.242	0.284	0.394
Years of education	0.208	0.018	< 2e-16
Ration card			
Above poverty line	Ref		
Below poverty line	-0.713	0.059	< 2e-16
No ration card	-0.379	0.073	0.000
Village level variance	0.54		
Community level variance	0.17		
AIC	13808		
LL	-6884		

Table SA 4.14: Results of logistic regression analysis showing effect of proportion of village with private electricity on individual latrine ownership

Fixed Part	Beta	SE	P
Intercept	-2.885	0.282	< 2e-16
Proportion of village with private electricity	0.195	0.700	0.780
Proportion of contacts with a latrine	0.807	0.082	< 2e-16
Caste			
Scheduled caste	Ref		
Scheduled tribe	0.026	0.131	0.845
OBC	0.659	0.069	< 2e-16
General	0.848	0.095	< 2e-16
HH electricity			
Private	Ref		
Govt.	-0.957	0.063	< 2e-16
No	-1.467	0.165	< 2e-16
# of beds in house	0.190	0.018	< 2e-16
# of rooms in house	0.349	0.019	< 2e-16
Roof-type			
Thatch	Ref		
Tile	-0.318	0.256	0.214
Stone	0.533	0.258	0.039
Sheet	0.467	0.258	0.070
RCC	1.439	0.261	0.000
Other	-0.242	0.284	0.394
Years of education	0.208	0.018	< 2e-16
Ration card			
Above poverty line	Ref		
Below poverty line	-0.713	0.059	< 2e-16
No ration card	-0.379	0.073	0.000
Village level variance	0.54		
Community level variance	0.17		
AIC	13808		
LL	-6884		

Table SA 4.15: Results of logistic regression analysis showing effect of village mean number of beds within household on individual latrine ownership

Fixed Part	Beta	SE	P
Intercept	-2.884	0.282	< 2e-16
Village mean number of beds in household	0.117	0.159	0.461
Proportion of contacts with a latrine	0.805	0.082	< 2e-16
Caste			
Scheduled caste	Ref		
Scheduled tribe	0.027	0.131	0.837
OBC	0.660	0.069	< 2e-16
General	0.848	0.095	< 2e-16
HH electricity			
Private	Ref		
Govt.	-0.957	0.063	< 2e-16
No	-1.467	0.165	< 2e-16
# of beds in house	0.189	0.018	< 2e-16
# of rooms in house	0.349	0.019	< 2e-16
Roof-type			
Thatch	Ref		
Tile	-0.318	0.256	0.213
Stone	0.531	0.258	0.039
Sheet	0.467	0.258	0.070
RCC	1.440	0.261	0.000
Other	-0.242	0.284	0.394
Years of education	0.208	0.018	< 2e-16
Ration card			
Above poverty line	Ref		
Below poverty line	-0.714	0.059	< 2e-16
No ration card	-0.380	0.073	0.000
Village level variance	0.54		
Community level variance	0.17		
AIC	13808		
LL	-6884		

Table SA 4.16: Results of logistic regression analysis showing effect of village mean number of rooms within household on individual latrine ownership

Fixed Part	Beta	SE	P
Intercept	-2.891	0.281	< 2e-16
Village mean number of rooms in household	-0.283	0.242	0.241
Proportion of contacts with a latrine	0.811	0.082	< 2e-16
Caste			
Scheduled caste	Ref		
Scheduled tribe	0.025	0.131	0.851
OBC	0.657	0.069	< 2e-16
General	0.847	0.095	< 2e-16
HH electricity			
Private	Ref		
Govt.	-0.958	0.063	< 2e-16
No	-1.466	0.165	< 2e-16
# of beds in house	0.190	0.018	< 2e-16
# of rooms in house	0.348	0.019	< 2e-16
Roof-type			
Thatch	Ref		
Tile	-0.317	0.256	0.215
Stone	0.532	0.258	0.039
Sheet	0.467	0.258	0.070
RCC	1.439	0.261	0.000
Other	-0.245	0.284	0.388
Years of education	0.208	0.018	< 2e-16
Ration card			
Above poverty line	Ref		
Below poverty line	-0.712	0.059	< 2e-16
No ration card	-0.379	0.073	0.000
Village level variance	0.53		
Community level variance	0.17		
AIC	13807		
LL	-6883		

Table SA 4.17: Results of logistic regression analysis showing effect of proportion of village that is Hindu on individual latrine ownership

Fixed Part	Beta	SE	P
Intercept	-2.882	0.281	< 2e-16
Proportion of village that is Hindu	-2.083	1.057	0.049
Proportion of contacts with a latrine	0.806	0.082	< 2e-16
Caste			
Scheduled caste	Ref		
Scheduled tribe	0.023	0.131	0.859
OBC	0.660	0.069	< 2e-16
General	0.849	0.095	< 2e-16
HH electricity			
Private	Ref		
Govt.	-0.957	0.063	< 2e-16
No	-1.466	0.165	< 2e-16
# of beds in house	0.190	0.018	< 2e-16
# of rooms in house	0.348	0.019	< 2e-16
Roof-type			
Thatch	Ref		
Tile	-0.318	0.256	0.214
Stone	0.529	0.258	0.040
Sheet	0.466	0.258	0.071
RCC	1.435	0.261	0.000
Other	-0.242	0.284	0.394
Years of education	0.208	0.018	< 2e-16
Ration card			
Above poverty line	Ref		
Below poverty line	-0.713	0.059	< 2e-16
No ration card	-0.380	0.073	0.000
Village level variance	0.50		
Community level variance	0.17		
AIC	13804		
LL	-6882		

Table SA 4.18: Results of logistic regression analysis showing effect of mean age respondents in village on individual latrine ownership

Fixed Part	Beta	SE	P
Intercept	-2.887	0.281	< 2e-16
Village respondent mean age	-0.169	0.066	0.011
Proportion of contacts with a latrine	0.805	0.082	< 2e-16
Caste			
Scheduled caste	Ref		
Scheduled tribe	0.027	0.131	0.834
OBC	0.663	0.069	< 2e-16
General	0.854	0.095	< 2e-16
HH electricity			
Private	Ref		
Govt.	-0.959	0.063	< 2e-16
No	-1.464	0.165	< 2e-16
# of beds in house	0.191	0.018	< 2e-16
# of rooms in house	0.349	0.019	< 2e-16
Roof-type			
Thatch	Ref		
Tile	-0.316	0.256	0.216
Stone	0.526	0.258	0.041
Sheet	0.465	0.258	0.071
RCC	1.432	0.261	0.000
Other	-0.244	0.284	0.391
Years of education	0.211	0.018	< 2e-16
Ration card			
Above poverty line	Ref		
Below poverty line	-0.711	0.059	< 2e-16
No ration card	-0.376	0.073	0.000
Village level variance	0.53		
Community level variance	0.17		
AIC	13802		
LL	-6881		

Table SA4.19 Results of multi-level logistic regression analysis showing the effect of social group latrine ownership proportions on individual latrine ownership at the levels of the individual, the community, and the village

	Model 1 Individual attributes	Model 2 Model 1 plus social contact latrine	Model 3 Model 2 plus village and community characteristics	Model 4 Model 3 plus community and village latrine ownership	Model 5 community latrine plus village level fixed effects (not shown)
Fixed Part	Beta (SE)	Beta (SE)	Beta (SE)	Beta (SE)	Beta (SE)
Intercept	-2.75 (0.29)	-2.9 (0.28)	-2.885 (0.28)	-2.773 (0.264)	-2.337*** (0.324)
Caste					
Scheduled caste	Ref	Ref	Ref	Ref	Ref
Scheduled tribe	0.08 (0.13)	0.03 (0.13)	0.016 (0.131)	-0.065 (0.126)	0.033 (0.129)
OBC	0.8*** (0.07)	0.66*** (0.07)	0.658*** (0.069)	0.591 (0.066)	0.625*** (0.067)
General	1.04*** (0.09)	0.85 *** (0.09)	0.842*** (0.095)	0.766 (0.09)	0.854*** (0.094)
HH electricity					
Private	Ref	Ref	Ref	Ref	Ref
Govt.	-0.97*** (0.06)	-0.96 *** (0.06)	-0.956*** (0.063)	-0.923 (0.061)	-0.939*** (0.062)
No	-1.49*** (0.16)	-1.47*** (0.16)	-1.464*** (0.164)	-1.433 (0.161)	-1.428*** (0.161)
# of rooms in house	0.35*** (0.02)	0.19*** (0.02)	0.348*** (0.019)	0.348 (0.019)	0.347*** (0.019)
# of beds in house	0.19*** (0.02)	0.35*** (0.02)	0.191 (0.018)	0.185 (0.018)	0.186*** (0.018)
Roof-type					
Thatch	Ref	Ref	Ref	Ref	Ref
Tile	-0.33 (0.26)	-0.32 (0.26)	-0.316 (0.256)	-0.244 (0.251)	-0.288 (0.253)
Stone	0.56* (0.26)	0.53* (0.26)	0.523* (0.258)	0.486 (0.252)	0.544* (0.256)
Sheet	0.48 (0.26)	0.47 (0.26)	0.462 (0.258)	0.423 (0.252)	0.466 (0.255)
RCC	1.47*** (0.26)	1.44*** (0.26)	1.424*** (0.261)	1.348 (0.255)	1.419*** (0.259)
Other	-0.24 (0.28)	-0.25 (0.28)	-0.244 (0.284)	-0.258 (0.277)	-0.197 (0.281)
Years of education	0.06 *** (0.00)	0.06 *** (0.01)	0.214*** (0.018)	0.204 (0.018)	0.206*** (0.018)
Ration card					
Above poverty line	Ref	Ref	Ref	Ref	Ref
Below poverty line	-0.75*** (0.06)	-0.71*** (0.06)	-0.71*** (0.059)	-0.698 (0.057)	-0.708*** (0.059)
No ration card	-0.41*** (0.07)	-0.38*** (0.07)	-0.372*** (0.073)	-0.366 (0.071)	-0.371*** (0.072)

Table SA4.19 continued

	Model 1 Individual attributes	Model 2 Model 1 plus social contact latrine	Model 3 Model 2 plus village and community characteristics	Model 4 Model 3 plus community and village latrine ownership	Model 5 community latrine plus village level fixed effects (not shown)
Proportion of contacts with a latrine		**0.81*** (0.08)**	**0.806*** (0.082)**	**0.823 (0.081)**	**0.743*** (0.081)**
Mean age in village			-0.124 (0.067)	-0.036 (0.033)	*
Proportion of village that is Hindu			-1.116 (1.212)	0.306 (0.66)	*
Community proportion Hindu			-0.669 (0.611)	-0.334 (0.437)	-0.53 (0.443)
Community mean number of rooms			-0.176 (0.087)	-0.418 (0.062)	-0.324*** (0.073)
Community mean education			0.084* (0.036)	-0.049 (0.027)	0.007 (0.032)
Mean age of community respondents			-0.015 (0.017)	-0.027 (0.013)	-0.021 (0.013)
Community latrine proportion				**2.58 (0.245)**	**1.782*** (0.258)**
Village latrine proportion				**1.683 (0.422)**	
Village level variance	0.72	0.54	0.48	.08	
Community level variance	0.20	0.17	0.16	>.001	>.001
AIC 13893		13806	13799	13769	13695
LL -6929		-6884	-6874	-6858	-6750

CHAPTER 5

Overview of our findings

Research in health behavior is increasingly showing the limitations of individual focused behavior change approaches and is instead focusing on the relationship between the individual and their social environment. Social network analysis adds critical insights to these perspectives. A network-based approach may be particularly critical in developing world contexts, where social connections, and the norms that are held within those connections, exert a particularly strong influence over individuals.

The aim of this dissertation was to investigate the different dimensions of social network effects on health within diverse settings. The first paper "Parenting influences on adolescent substance abuse outcomes" demonstrates that social network effects may work not only through homogenous peer-to-peer pathways, but through the more layered pathway of a peer's-parents to peer. The second paper "Social network correlates of latrine ownership in rural India" provides evidence that not only are an individual's direct social relations potentially influential to an individual's latrine ownership decision, but that the individual's structurally related network characteristics may impact that possible effect. Finally the third paper, "New perspectives on latrine ownership in India: the effect of social network communities" illuminates the potential utility of identifying network derived communities in order to understand the normative dynamics of individual latrine ownership decisions.

Applications to global health

While the social context of chapter 2 was adolescent social networks in the United States, the potential application of the outcome, that network effects transcend homogenous network relations, is global. This perspective may be particularly

important for looking at relationships in more traditional communities where complex social hierarchies significantly influence individual behavior. For instance, while we know qualitatively that in India mother-in-laws are very influential over the child rearing behavior of young Indian mothers, we can use methodology to expand our understanding of that dynamic by 1) linking mother-in-laws and daughters-in-laws to quantitatively correlate their attitudes and behaviors, and 2) test for the ways in which network related social influence at the mother-in-law level impacts the behaviors of daughter-in-laws and visa versa. For instance we may find that daughters-in-laws who are exposed to an educational campaign may influence each other's behavior, not through the direct channel of daughter-in-law to daughter-in-law but indirectly through the network of their mother-in-laws.

Analyses of this sort are best served through use of longitudinal data, however, which is currently scarce on diverse populations in developing countries. Future research should focus on the collection of sociocentric longitudinal data in different social and cultural contexts around the world. With access to richer and more culturally representative data, we can strengthen our understanding of these complex network processes and their potential application to interventions in a multitude of settings.

Chapters 3 and 4 explore a range of possible network effects within the same population in exploration of the same outcome. Using rich network data from 75 villages in rural India, we explore the outcome of household latrine ownership. Overall, results from both papers suggest that, consistent with previous non-network based research, social dynamics play an important role in latrine ownership decisions. Looking more deeply, several important aspects of the research are worth noting.

First, there is a strong correlation between the latrine ownership of an individual's social contacts, and the latrine ownership of the individual. This result was consistent whether we looked at each respondent and each of their friends individually, or whether we simply assigned a "social contacts latrine proportion" measure to each individual. While the lack of longitudinal data in this case makes it difficult to infer influence, there is strong evidence of a social dynamic at play. Further research could elucidate more exactly the nature of that dynamic.

For instance, we may discover that the correlation in latrine ownership between connected individuals is due to *homophily,* or the tendency for similar others to become socially connected. If in fact it is the similarities between connected individuals that determine both relationship formation and latrine adoption, we could potentially pinpoint those similarities and use them to create more customized interventions. If on the other hand, the correlation in latrine ownership is in fact due to some sort of social influence, interventionists can potentially leverage that social influence dynamic through innovations which are implemented through network channels.

Secondly, network position matters. Those more central to the network are more likely to own latrines, while those more peripheral are less likely. However, the correlation between latrine ownership of the individual and that of her social contacts is smaller for those more central. These results add support to the possibility that the latrine ownership correlation between individuals may be a matter of influence. It is more difficult to explain those results as due to homophily. We would have to assume that there is some other shared characteristic of peripheral people that simultaneously reduces their probability of latrine ownership while increasing their probability of being connected if they do happen to own a latrine. These results

suggest that while initially interventions may be more efficacious for those more central, once the central innovators have adopted, an approach directed towards the more peripheral could be warranted.

Finally, community level effects are stronger than the effect of either direct social contacts, or the effect of those in one's village. This provides evidence for community based norms that are held within socially relevant reference groups, and not simply based upon geographic proximity. This could have enormous potential for future intervention work, as the use of network-based communities could offer interventionists a powerful and so far underutilized tool for implementing behavior change. It also suggests that interventionists not using a network approach could be undermining their own efforts by neglecting a potentially powerful determinant of individual health behavior.

Future directions

Overall the results of this body of work strengthen our understanding of network effects on diverse health outcomes within 2 culturally, economically, and socially very different contexts. While previous research has highlighted the utility of incorporating a network perspective into health and development interventions, the results of this body of work can inform novel ways of doing so. Future research is warranted to further test the results of these three studies, and to develop effective ways to integrate network methods into the hand-on work being done by health promotion workers out in the field.