Title
Essays About Group Dynamics and Social Networks

Author
Gee, Laura Katherine

Publication Date
2013

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA, SAN DIEGO

Essays About Group Dynamics and Social Networks

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

in

Economics

by

Laura Katherine Gee

Committee in charge:

Professor James Andreoni, Chair
Professor Julie Cullen
Professor Gordon Dahl
Professor James Fowler
Professor Uri Gneezy

2013

The dissertation of Laura Katherine Gee is approved, and it is acceptable in quality and form for publication on microfilm and electronically:

Chair

University of California, San Diego

2013

DEDICATION

To my family and friends for supporting me during this process. In particular my sister Elizabeth for letting me live with her the first two years I was in San Diego. Also, a special thanks to my husband, Dan, for moving to San Diego to support me while I took on this endeavor.

EPIGRAPH

Enthusiasm is the mother of effort,
and without it nothing great was ever achieved.
—Ralph Waldo Emerson

TABLE OF CONTENTS

Signature Page iii

Dedication iv

Epigraph v

Table of Contents vi

List of Figures viii

List of Tables ix

Acknowledgements xi

Vita xii

Abstract of the Dissertation xiii

Chapter 1 Social Networks and Labor Markets: How Strong Ties Relate to Job Transmission On Facebook's Social Network . 1
- 1.1 Introduction 2
- 1.2 Literature Review 4
- 1.3 Data 7
  - 1.3.1 Job Transmission 8
  - 1.3.2 Measuring Tie Strength 10
- 1.4 Analysis 14
  - 1.4.1 Descriptive Weak Ties Hypothesis 17
  - 1.4.2 Conditional Weak Ties Hypothesis 20
  - 1.4.3 Robustness Checks 34
- 1.5 Concluding Remarks 41
- 1.6 Acknowledgement 43
- 1.7 Chapter 1 Appendix 43
  - 1.7.1 Robustness Checks 45
  - 1.7.2 In Text Tables Details & Complimentary Tables 65

Chapter 2 Gun For Hire: Delegated Enforcement and Peer Punishment in Public Goods Provision 81
- 2.1 Introduction 82
- 2.2 Background 85
- 2.3 The Games 88

2.3.1 The Linear Public Goods (LPG) game . . . 88
2.3.2 The Gun Hired (GH) game 89
2.3.3 The Gun For Hire (G4H) game 91
2.3.4 The Peer-to-Peer (P2P) game 92
2.3.5 The Gun For Hire and Peer-to-Peer (G4H/P2P) game 94
2.4 Procedures 95
2.5 Results 97
2.5.1 Baseline: LPG 97
2.5.2 Starting From Vigilante Justice: Welfare in Period 11-20 following P2P in Periods 1-10 . 103
2.5.3 Does Delegated Enforcement Crowd Out Peer Punishment? 105
2.6 Concluding Remarks 108
2.7 Acknowledgement 111

Chapter 3 The Nature of Giving Time to Your Child's School . . . 112
3.1 Introduction 113
3.2 School Contribution: Private and Public Benefits . 114
3.3 Method 117
3.3.1 Data 117
3.3.2 Dependent Variables 118
3.3.3 Independent Variables 119
3.3.4 Data Analysis 120
3.3.5 Results 121
3.4 Conclusion 130
3.5 Acknowledgement 131

Bibliography 132

LIST OF FIGURES

Figure 1.1: Tag Example 12
Figure 1.2: Posting Example 12
Figure 1.3: Mutual Friends and Network Overlap 14
Figure 1.4: Job Transmission Examples 15
Figure 1.5: Distribution of Tags (including 0) 19
Figure 1.6: Employment Information Example 44
Figure 1.7: Distribution of Tags (excluding 0) 52
Figure 1.8: Non-parametric Plots 59

Figure 2.1: Contributions to the Public Good after LPG 99
Figure 2.2: Average Per Subject Net Earnings after LPG 101
Figure 2.3: Contributions to the Public Good after P2P 105
Figure 2.4: Average Per Subject Net Earnings after P2P 106
Figure 2.5: Average Per Subject Peer Punishment Costs after P2P in Periods 1-10 109

LIST OF TABLES

Table 1.1: User Summary Statistics (Primary Sub-sample) 22
Table 1.2: Friend-level Summary Statistics (Primary Sub-sample) . . 22
Table 1.3: Dyad-level Tie Strength Summary Statistics (Primary Sub-sample) 24
Table 1.4: Correlation Absolute Tie Strength (Primary Sub-sample) . 24
Table 1.5: Correlation Percentage Tie Strength (Alternative Sub-sample) 25
Table 1.6: Dyad-level Demographic Summary Statistics 28
Table 1.7: Linear Models Dependent Variable Job Transmission Absolute Tie Strength Measures One-by-One (Primary Sub-sample) 30
Table 1.8: Linear Models Dependent Variable Job Transmission Scaled Tie Strength Measures One-by-One (Alternative Sub-sample) 31
Table 1.9: Linear Models Dependent Variable Job Transmission All Tie Strength Measures 34
Table 1.10: Primary Sub-sample Weighting Example 44
Table 1.11: Sensitivity to Controls in Linear Models Dependent Variable Job Transmission 45
Table 1.12: Robustness Against Incidental Job Transmission 47
Table 1.13: Robustness Against Different Tie and Tie Strength Definitions 53
Table 1.14: Friends of Friends 57
Table 1.15: Linear Models Dependent Variable Job Transmission All Tie Strength Measures (Replication of Results in Table 1.9 in Text Using Other Sub-samples for Each Model) 60
Table 1.16: Conditional Logit Models Dependent Variable Job Transmission (Alternative Sub-sample) 63
Table 1.17: User statistics (Alternative Sub-sample) 65
Table 1.18: Friend statistics (Alternative Sub-sample) 66
Table 1.19: Dyad-level Tie Strength Summary Statistics (Alternative Sub-sample) 66
Table 1.20: Correlation Absolute Tie Strength (Alternative Sub-sample) 67
Table 1.21: Dyad-level Demographic Summary Statistics 67
Table 1.22: Absolute Tie Strength Measures One-by-One, Full Results of Table 1.7 in Text 68
Table 1.23: Linear Models Dependent Variable Job Transmission Scaled Tie Strength Measures One-by-One (Alternative Sample), Full Results of Table 1.8 in Text 71
Table 1.24: Scaled Tie Strength Measures One-by-One (Primary Sub-sample replication of Table 1.8 in Text) 74

Table 1.25: Correlation Percentage Tie Strength (Primary Sub-sample replication of Table 1.5 in Text) 77
Table 1.26: Linear Models Dependent Variable Job Transmission All Tie Strength Measures, Full Results of Table 1.9 in Text 78

Table 2.1: Equilibrium Predictions 95
Table 2.2: Average Earnings per Subject with baseline LPG 100
Table 2.3: Determinants of Earnings in Rounds 11-20 by treatment . . 102
Table 2.4: Average Earnings per Subject after P2P in Periods 1-10 . 105

Table 3.1: Dependent Variables Measures of Time Contribution (standard errors), abridged results 123
Table 3.2: Household Has A Say in Student's Placement (standard errors) 129

ACKNOWLEDGEMENTS

A special thanks to my dissertation committee members: Jim Andreoni, Julie Cullen, Gordon Dahl, James Fowler, and Uri Gneezy. Also a huge thanks to my co-author, Jason Jones. Additionally I would like to thank the anonymous reviewers, and the participants at the following seminars and conferences: UCSD , ESA 2010, SWET 2011, Pomona College, APPAM 2012, George Mason University ICES, WEAI 2012, Political Networks 2012, BABEEW 2012, SCCAM 2012, Clemson University, Bryn Mawr College, Tufts University, Quinnipiac University, USC Marshall, Queens College CUNY, Hofstra University, University of Toronto Rotman, and Santa Clara University. Additionally I would like to thank the following people for thier comments on this research: Kristy Buzard, Leah Lakdawala, Simon Gachter, Nikos Nikiforakis and Anya Savikhina.

Chapter 1, in part is currently being prepared for submission for publication of the material. Gee, Laura; Jones, Jason. The dissertation author was the primary investigator and author of this material.

Chapter 2, in full, is a reprint of the material as it appears in the Journal of Public Economics, vol. 96, no. 11-12, December 2012, p. 1036-1046. Gee, Laura; Andreoni, James., 2012. The dissertation/thesis author was the primary investigator and author of this paper.

Chapter 3, in full, is a reprint of the material as it appears in Nonprofit and Voluntary Sector Quarterly, vol. 40 no. 3 June 2011, p. 552-565. Gee, Laura, 2011. The dissertation/thesis author was the primary investigator and author of this paper.

VITA

2004	B.A. in Economics *cum laude*, Barnard College Columbia University
2009	M.A. in Economics, University of California, San Diego
2013	Ph.D. in Economics, University of California, San Diego

PUBLICATIONS

Gee, "The Nature of Giving Time to Your Child's School", *Nonprofit and Voluntary Sector Quarterly*, 2011.

Gee, Andreoni "Gun For Hire: Delegated Enforcement and Peer Punishment in Public Goods Provision", *Journal of Public Economics*, 2012

ABSTRACT OF THE DISSERTATION

Essays About Group Dynamics and Social Networks

by

Laura Katherine Gee

Doctor of Philosophy in Economics

University of California, San Diego, 2013

Professor James Andreoni, Chair

This dissertation looks at how individual decisions are affected by group dynamics and social networks in a variety of settings. The first chapter examines the effect of social networks in labor markets by specifically asking what type of friend is most likely to help an individual find a new job. The second chapter uses a lab experiment to compare two methods to encourage socially optimal provision of a public good. The third chapter looks at the specific public good of parental volunteering at a child's school and disentangles whether parents' choices to volunteer are motivated by public or private benefits.

Chapter 1

Social Networks and Labor Markets: How Strong Ties Relate to Job Transmission On Facebook's Social Network

1.1 Introduction

Social networks help explain observed labor market phenomenon like duration dependence and the socioeconomic, geographic, and racial concentration of unemployment. Social networks are essential to job search because over 50% of jobs are found through a social contact.[1] Additionally individuals who find a job via a social contact have longer tenure, higher income and higher productivity.[2] Thus, understanding how individuals use their social networks to obtain new employment is economically relevant. Perhaps the most widely-known finding in this literature is the "strength of weak ties," claim by Granovetter (1973, 1995), which has heavily influenced research in social networks and labor markets.[3] This hypothesis states that most jobs are found through weak ties rather than strong ties. A weak tie describes a social connection between a person and a friend (a dyad) who are not close to each other. Closeness is a general term, but some have defined it as amount of time spent together, number of mutual friends, or subjective reports of how a person feels about a friend. The strength of weak ties claim is important, because it asserts not all social ties are equally useful in job finding.

Using data from 296 million Facebook friendships, representing 6 million US users and their friends, we decompose the original strength of weak ties claim into two distinct hypotheses about what type of friend is most helpful in job finding. The Facebook data are more nationally representative than previous network data, have excellent information on the structure of an individual's Facebook friendships, and allow for more objective measures of contact. We use two types of tie strength measures: contact-based and

[1] See Topa (2011), Jackson (2011), Munshi (2011), Ioannides and Loury (2004) and Marsden and Gorman (2001) for an extensive review of the literature in Economics and Sociology

[2] See the literature review section for citations.

[3] Evidence of the influence of this paper can be seen in its very high level of citations. In September 2012 Granovetter (1973) had 21,181 citations on Google Scholar and 740 citations on JSTOR, in contrast another influential paper "The Market for Lemons" (Akerlof, 1970) published three years before had only 14,810 citations on Google Scholar and 601 citations on JSTOR.

structure-based, both of which have been popular in the theoretical and empirical literature. We are able to test whether contact-based measures are a good proxy for structure-based measures, and this is the first paper that the authors are aware of that uses both types of tie strength measures to explain labor market outcomes. Specifically, the data include two types of online contact, tags and posts, as the contact-based measures of tie strength and the number of mutual friends as the network structure-based measure of tie strength. We measure help in job search by recording job "transmission." A job transmission occurs when a person and a friend (a) both work at the same employer (b) became Facebook friends at least one year before the person's most recent start date and (c) the friend began working at the shared employer at least a year before the other person's start date. Because of the specific timing of events a job can only be transmitted from a user to a friend or from a friend to a user, but not both. Using this definition, 7% of Facebook users were transmitted their job through a friend.[4] We test the relationship between job transmission and tie strength by testing two distinct hypotheses.

The first hypothesis is meant to closely mirror the analysis in the seminal paper (Granovetter, 1973), which we refer to as the "Descriptive Weak Ties Hypothesis." The Descriptive Weak Ties Hypothesis asserts that most jobs are found through a weak tie rather than a strong tie when concentrating on the relationship between a person and the specific friend(s) who helped. Our data support the Descriptive Weak Ties Hypothesis, as over 90% of individuals who find a job do so through a very weak tie. However, the distribution of tie strength in the population at large is also highly skewed toward weak ties, so the strength of weak ties is that individuals have so many weak ties.

In addition, we introduce the "Conditional Weak Ties Hypothesis" which states that the probability of being helped by a particular friend is

[4]Our data include many job transmission at retail and government employers. Retail and the government made up 7.9% and 13% of US GDP in 2009 (http://www.nrf.com/modules.php?name=Pages&sp_id=1214). We believe this is a large and generally understudied portion of the US labor market.

decreasing in tie strength conditional on information about the individual's whole network. When controlling for individual-level heterogeneity using a user-level fixed effect and myriad dyad controls to mitigate dyad-level heterogeneity, we find that *increases* in tie strength are associated with *increases* in the probability of job transmission from a particular friend. This positive relationship is evidence against the Conditional Weak Ties Hypothesis. For example, a dyad with at least one contact-based interaction in the past year has almost twice the probability of job transmission as a dyad with no contact over the same time period.[5]

We find support for the Descriptive Weak Ties Hypothesis but reject the Conditional Weak Ties Hypothesis. In short, a person is most likely to be transmitted a job from a weak tie because weak ties are prolific in social networks. However, when taking into account information about all of an individual's social ties, a strong tie has a higher probability of job transmission. These findings inform future theoretical modeling through emphasizing the importance of the information used in the model and differences in how tie strength is measured. Additionally, these results suggest that targeted encouragement of both formal and informal social networking might increase job finding.

1.2 Literature Review

Previous empirical work has shown that networks can influence labor market outcomes. In most cases, the effect is positive: less unemployment, higher income, higher productivity, and longer job tenure (Brown et al., 2012; Beaman, 2012; Beaman and Magruder, 2012; Cappellari and Tatsiramos, 2010; Mayer, 2012; Shue, 2012; Wei et al., 2012; Schmutte, 2010; Bandiera et al., 2009; Babcock, 2008; Tassier, 2006; Loury, 2006; Castilla, 2005; Elliott, 1999;

[5]Contact is defined as a tag or post on Facebook. These terms are described in more detail later in the text.

Marmaros and Sacerdote, 2002; Topa, 2001; Simon and Warner, 1992). Understanding how individuals use their social networks has been of both empirical and theoretical interest. In general, theoretical models define tie strength as a binary measure of existence of the relationship (Bramoulle and Saint-Paul, 2010; Calvo-Armengol et al., 2007; Calvo-Armengol and Jackson, 2004; Ioannides and Soetevent, 2006) or make assumptions about ties of different strengths (Zenou, 2011; Montgomery, 1992; Boorman, 1975). Additionally, some models have shown the positive attributes of jobs found through referrals (Galenianos, 2011; Montgomery, 1991). This analysis finds that both the structure and amount of contact between a social tie are related to labor market outcomes, and so hopefully will inspire further theoretical models with this in mind.

Previous studies concentrating specifically on the relationship between tie strength and job search have often tested the Descriptive Weak Ties Hypothesis, which does not take into account information about the whole social network (usually because the information was unavailable). When using this type of analysis, Granovetter (1973) found that most jobs are discovered through a weak tie. In his theoretical model, Granovetter defines strong ties as occurring between those dyads who share a larger proportion of mutual friends, and that a strong tie is never the only path between two nodes (a bridge). Because strong ties are never a bridge, weak ties convey more novel and useful job information in his theoretical model. Although Granovetter's theory uses a network structure-based measure of tie strength, in his empirical analysis, he uses a contact-based measure of tie strength: self-reported amount of contact between friends. He reports the results of a survey of 54 recent job-changers who obtained their current job through the help of a social tie. He asked these individuals how often they saw that friend around the time that she passed on the job information.[6] Of those who found their most recent job from a

[6] He defines "often" as at least twice a week, "occasionally" as more than once a year but less than twice a week, and "rarely" as once a year or less.

friend, 26.7% saw that person "often," 55.6% saw that person "occasionally," and 27.8% saw the person "rarely." So, Granovetter concludes that most jobs are found through a weak tie. Even though Granovetter's strength of weak ties finding has not always been replicable it has had significant influence on the literature about social networks and job search.[7]

The theoretical foundation for the strength of weak ties conjecture is that weak ties offer better access to useful information because they bridge across structural holes (Burt, 2004). Some studies on information diffusion in social networks have supported the efficacy of weak ties in information diffusion (Grabowicz et al., 2012; Bakshy et al., 2012; Ugander et al., 2012; Lin et al., 1978), while others have not (Harrigan et al., 2012; van der Leij and Goyal, 2011; Friedkin, 1980). The type of tie most useful for information diffusion differs based on the type of social network, the definition of tie strength, and the information being transmitted. It is likely that contact-based measures of tie strength measure the bandwidth of communication between two individuals, while network structure-based measures (like number of mutual friends) act as a proxy for the diversity of information. There is evidence of a trade-off between novelty of information and rate of information transmission between two individuals (Cui et al., 2012; Aral and Alstyne, 2011; Onnela et al., 2007; Reagans and McEvily, 2003).[8] Consider the following extreme example. An individual's strongest tie is to herself. Yet she has no new or useful information to offer herself, even if she is constantly in contact with herself. At the other extreme, the weakest tie is to a total stranger. A stranger may have access to lots of novel and useful information, but because there is no contact between an individual and a stranger there is no means for that information to reach its target.

[7]For example Bian (1997) found strong ties matter most while Yakubovich (2005) confirmed the weak ties finding, see Lin (1999) for a review. Evidence of the influence of Granovetter (1973) can be seen from its high citation rate as detailed in footnote 3.

[8]There is also a related line of research exploring a tradeoff between type of favor and network density (Karlan et al., 2009).

It is intuitively appealing that weak ties are most useful for job search because they offer access to novel information. Still, it seems equally intuitive that strong ties might be most useful because they are in contact more often, have superior information about each other, more reputation at stake for giving a referral, and more opportunities for the favor to be returned at a future date. A few studies imply that strong ties matter for employment. Studies using information about an individual's network of schoolmates, co-workers, or neighbors have found support for the efficacy of these strong ties in job finding (Rider, 2011; Bayer et al., 2008). Bayer et al. (2008) is especially compelling because they show the causal, rather than correlational, effect of being neighbors on the likelihood of working together. Finding the causal effect of an individual's social ties on an outcome for that individual can be difficult, and so we detail how our identification strategy relates to the peer effects literature in the analysis section of this paper (Sacerdote, 2011; Manski, 1993).

1.3 Data

The data include information about individuals and their friends from the social networking website Facebook. Facebook has over 1 billion active users, and an estimated 167 million users in the US.[9] In the US, over 54% of adults have a Facebook account, and 40% of social network users have "friended" their closest friends on social networking websites (Burke and Kraut, 2012; Bakshy et al., 2012). Understandably, the network that a user has on Facebook is not an exact representation of her true network, and a large amount of unobservable contact takes place outside of Facebook. We acknowledge these shortcomings of our data. Still, previous work has shown that Facebook interaction is correlated with other interactions (Jones et al.,

[9]See http://newsroom.fb.com/imagelibrary/downloadmedia.ashx?MediaDetailsID=4227&SizeId=-1 and http://www.socialbakers.com/facebook-statistics/?interval=last-6-months#chart-intervals

2012). We restrict our analysis to users and friends from the US between the age of 16 to 64, with education and employer information, and who have been on Facebook for at least 365 days before the user of interest's most recent start date at her currently listed employer. These restrictions result in a sample of 6 million users and their friends, or 296 million dyads.[10]

1.3.1 Job Transmission

Ideally, we would directly observe if an individual obtained her most recent job with the help of a friend, but instead we measure job "transmission." Users on Facebook can report their employment history, including their current and past employers, position, start date, and end date for each position listed.[11] We define a user as being "transmitted" her most recent job from a friend if the following criteria are met:

1. The user and this friend worked at the same employer.[12]
2. The user began working at that employer at least 365 days after their friend started working at that employer.
3. The user and the friend were friends on Facebook at least 365 days before the user started working at that employer.

The requirement that a user (ego) began working at the employer at least 365 days after her friend (alter) should exclude cases in which an ego and alter

[10]The data exclude people who do not have a valid start date listed for their employment and those who have listed their employer as "self-employed," "unemployed," or as "stay-at-home parent." Also, the data omit users who have two jobs that started at exactly the same time, so there is only one observation for each dyad. For a discussion of how these individuals differ from the Facebook population as a whole and the US population see footnote 36

[11]See the Appendix Figure 1.6 for a picture of how this information is recorded in Facebook.

[12]Users self-report their employer's name, and Facebook assigns each of these self-reported names a unique employer ID number. We use these unique employer ID numbers to identify if a user and friend share an employer. It is possible for variations on the same employer name to be given different employer ID numbers, and in this case we may not identify if a user and friend work together. This procedure may understate the number of job transmissions.

jointly apply at the same employer but the alter starts employment at that employer slightly before the ego. Also, it is more likely that an alter with a year long tenure at an employer would be able to help an ego find a job. The requirement that the alter and ego have been friends on Facebook for at least a year before they work together ensures there is at least one year's worth of Facebook interactions for the dyad. If interviews take place less than one year before an ego's start date, this restriction has the added effect of excluding cases in which an ego becomes friends with an alter during the interview process.[13]

Our measure of job transmission is not a perfect analog to those definitions used in previous survey work. An interesting feature of our job transmission definition is that multiple friends can help a user find her most recent job, while previous work has concentrated on the most helpful social connection. In reality, however, many friends may help. Our definition of job transmission is more restrictive in many ways than self-reported help because it requires that the dyad work together and that the friend have tenure at that employer of at least one year. However, this measure may also attribute job transmission when an individual would not self-report a friend as being helpful. Subjective measures of job helping are likely to have systematic errors, leading to biased coefficients. While our measure also suffers from measurement error, we have

[13]For example, imagine a user who interviews for a job at UC San Diego in December 2011 and becomes Facebook friends with a UC San Diego faculty member she meets during that interview process during December 2011. If the user begins her job in September 2012, less than a year after becoming friends with the faculty member, then this would not be measured as a job transmission. In short, if interviews occur less than one year before a user starts her job, then we will not measure friendships which result from an interview as job transmission friendships. Many of the employers in our data set employers that tend to hire individuals quite close to their start date (e.g. retailers). So there is no reason to believe that these employers generally interview employees over a year prior to hiring. There is reason to believe that most interviews that take place well in advance are for jobs that require a college education but analysis using only college educated users and their friends are of the same sign and significance as the results reported in text (see the Robustness Checks section).

have no reason to suspect these errors are systematic.[14] Using this measure of job transmission, about 7% of the 6 million Facebook users in the sample were transmitted their most recent job from a friend. This level is well below the 50% found by previous works (Topa, 2011; Granovetter, 1973), and the difference may be largely due to our highly restrictive job transmission definition.[15]

1.3.2 Measuring Tie Strength

In general, a weak tie is a friendship where the individual does not feel very close to a particular social connection. This paper concentrates on two of the most widely used types of tie strength measures: contact-based tie strength and network structure-based tie strength. To most closely match the measures used by Granovetter (1973), contact is measured between a user and a friend for the full year before the user started her most recent job, and network structure is measured using the number of mutual friends shared by a user and a friend a year before the user's job starts. All of these tie strength measures were computed using only fully anonymous data without any personal identifiers such as first or last name.

The contact-based measures of tie strength used are photo tags and wall posts on Facebook. A photo tag, as pictured in Figure 1.1, is observed when a

[14] We are currently in the process of running a survey with Facebook to obtain subjective measures of job help, which we plan to use in future research.

[15] To benchmark what portion of this 7% might be due to the random chance that all three our job transmission criteria are met, we did a permutation test. Each user has an employer/start-date pair. We randomly re-assigned these employer/start-date pairs without replacement to other users in the data and then checked if a job transmission still occurred for each user. Because this is computationally intensive we began with a random sample of 27,000 dyads representing 120 users with an actual job transmission rate of 6.2% for this sub-sample. We did the permutation 1,000 times and found that the average rate of job transmission was only 0.3%. As an added check we did the same procedure but restricted employer/start-date pairs to be randomly re-assigned within the same city. In this second permutation test a user who was employed by Walmart in Oakland, California in September 2010 could only be randomly re-assigned another job that was from Oakland, California. Under this test the random rate of transmission was still only 1.5% which is well below the 7% we observe in the actual data.

user marks a photo with a friend's name so that the friend and other Facebook users may more easily locate the photograph. Photo tags may be seen as evidence of real world interaction.[16] A wall post, as pictured in Figure 1.2, is observed when a user posts a message on the Facebook homepage (wall) of a friend. Although there are many measures of contact available for Facebook users this analysis concentrates on tags and wall posts because these measures have been shown to be good predictors of real world friendships (Jones et al., 2012).[17] We exclude more commonly used modes of contact, which on a social network like Facebook may represent incidental rather than meaningful contact. In addition to the raw number of tags and posts, we also use the number of tags or posts as a percentage of total tags and posts by the user in the previous year.[18] These measures of contact-based tie strength are most similar to those used in previous research to test the Descriptive Weak Ties Hypothesis.

The network structure-based measure of tie strength is the number of mutual friends shared by a user and a friend. A scaled measure of mutual friends as a percentage of possible mutual friends, network overlap, is also included in the analysis.[19] Network overlap is defined as $O_{ik} = \frac{m_{ik}}{d_i - 1 + d_k - 1 - m_{ik}}$, where m_{ik} is the number of mutual friends between i and k, d_i is the number

[16]For example, if user A and friend B are together and A takes a photo of B, then it is likely that A will tag B in the photo when she uploads the photo to Facebook. Thus a tag from A to B is a proxy for real world interaction.

[17]At the mean start date for the 6 million US users in our sample the modes of communication from most to least used were: Comment, Like, Message, Wall Post, Tag, Poke and Chat.

[18]For example, consider user A who has two friends named B and C. If user A tagged person B two times last year and she tagged person C six times last year, we would say that 25% ($\frac{2}{2+6}$) of user A's tags were to person B and that the remaining 75% ($\frac{6}{2+6}$) was to person C. We compute this with a denominator based on the friends we observe in our data, rather than all Facebook friends. In practice that means that percentage tags and posts are estimates for the primary sub-sample. Other reasons that we do not include all Facebook friendships are that not all dyads have been friends for at least one year, or not all dyads are eligible for a job transmission (e.g. one resides outside the US).

[19]This is also known as a Jaccard Index, and is related to the concept of support introduced by Jackson et al. (Forthcoming). Support measures if any mutual friends exist.

Figure 1.1: Tag Example

Note: This figure shows a photo tag, which is observed when a user indicates that another user is pictured in a photograph. The number of times a user tags a friend in a year is one of our contact-based tie strength measures.

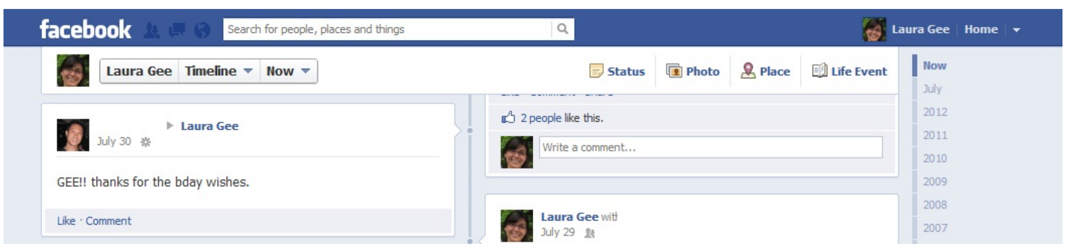


Figure 1.2: Posting Example

Note: This figure shows a post, which is a message displayed on a Facebook user's homepage (wall). The number of times a user posts on a friend's wall in a year is one of our measures of contact-based tie-strength measures.

of friends of person of i (degree of i), and d_k is defined analogously.[20] Then $O_{ik} = 1$ if i and k have all the same friends, and $O_{ik} = 0$ if they have no friends in common. This measure of overlap was chosen because it closely mirrors the spirit of Granovetter's theoretical model.[21] Figure 1.3 shows an example of

[20]We treat the Facebook graph as un-directed meaning that if a tie (edge) exists from i to j, then a tie also exists from j to i.

[21]Granovetter (1973) states "the stronger the tie between A and B, the larger proportion of individuals S to whom they will both be tied, that is, connected by a weak or strong

how network overlap is computed. Figure 1.4 shows two examples of the order of events that would identify a dyad as having a job transmission and the time frame used to measure contact-based and structure-based tie strength for these dyads.

As mentioned in the literature review, previous studies have found a diversity-bandwidth tradeoff, meaning that more novel/diverse information tends to come from friends who have a lower rate of contact. Dyads with fewer mutual friends or less network overlap may convey more novel information between them and thus are more likely to have a job transmission. If one looks at the network pictured in the far left portion of Figure 1.3, person A and B have no mutual friends, so if person D or E has information about a job opening of interest to A, then that novel information would have to flow through B. However, in the network with 1 mutual friend (network overlap = 1/3) we see the information from D or E can now come through three different paths. The more mutual friends there are, the more paths there are for the information to flow, but the novelty of information may be lower. Network structure-based measures may approximate the novelty of information, while contact-based measures may approximate the rate of information flow. So contact-based measures may not be a good proxy for structure-based measures.

Both our contact-based and network structure-based measures are dyad-level measures. This enables us to use a random sub-sample of the complete network instead of using the whole graph or a component of the total graph.[22] If the complete network, meaning all individuals and all their ties, was used, then any assumption of independence across observations would likely be violated. Using a random sub-sample of individuals and all their friends makes the assumption of independence across dyads more convincing.[23] In the analysis,

tie. This overlap in their friendship circles is predicated to be least when their tie is absent, most when it is strong, and intermediate when it is weak."

[22]A component is a maximal subnetwork where every pair of nodes is connected by some sequence of links.

[23]In our data each dyad only occurs a single time. For example, imagine all Facebook users have a unique ID number. Then we observe the dyad with user 1 and friend 10 only

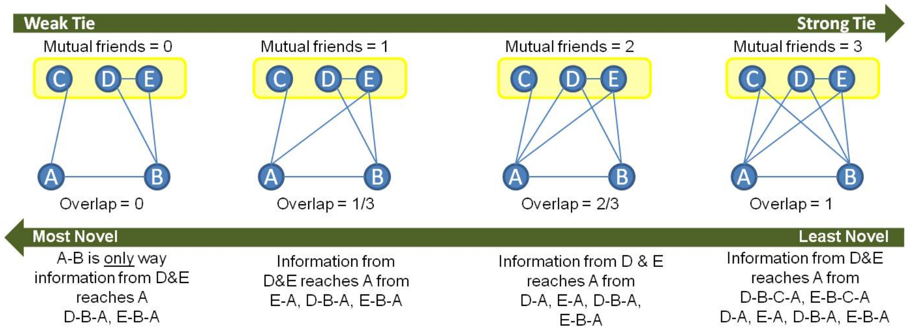


Figure 1.3: Mutual Friends and Network Overlap

Note: Network overlap is measured as the number of mutual friends as a percentage of possible mutual friends. Mutual friends and network overlap are the structure-based tie strength measures. Along the bottom of the figure is an illustration of how higher network overlap (stronger ties) may have less novel information flow between them.

the model includes an individual-level fixed effect, which precludes measures that do not vary within the individual, like the individual clustering coefficient, betweenness, or degree centrality for a given node i. The inclusion of a user-level fixed effect is essential for identification in the model, so the use of a dyad-level measure like mutual friends is also essential to the identification strategy.

1.4 Analysis

Exploring the causal relationship between tie strength and job transmission is difficult. First, the network itself is endogenously determined, and second the level of tie strength between each dyad in the network is also endogenously chosen. We are most interested in the question of how tie strength affects job transmission, so we will take the network as given.[24] Taking the

a single time. Meaning we see only the dyad with $i = 1, k = 10$ but not also the dyad containing $i = 10, k = 1$.

[24] We believe that the effect of network selection on job transmission is an important area for research. The previous work on exogenous assignments of networks has concentrated

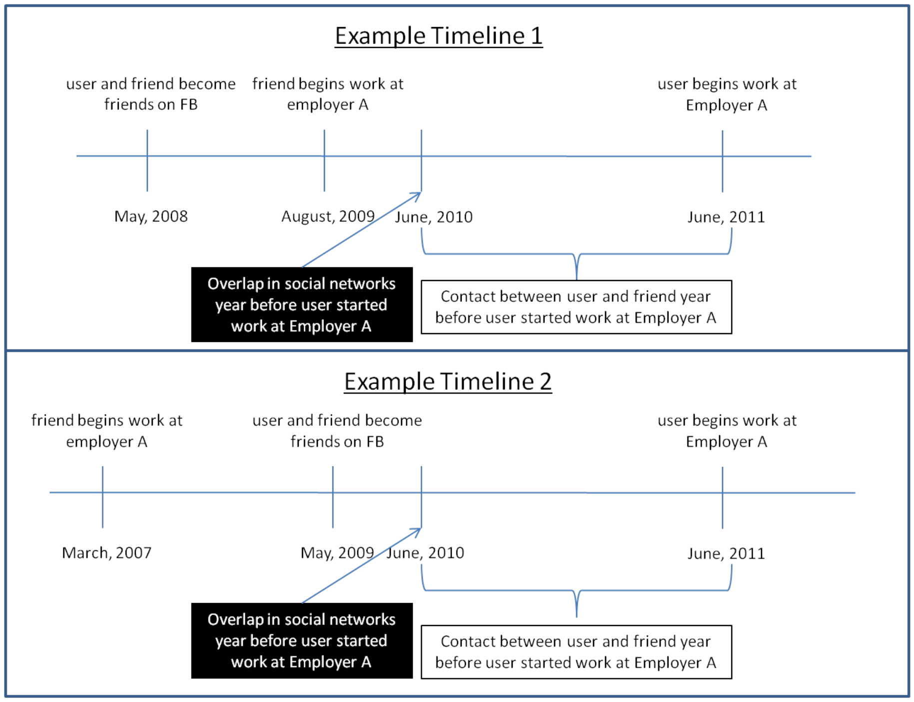


Figure 1.4: Job Transmission Examples

Note: This figure shows two examples of the timeline of events where a user would be identified as being "transmitted" her most recent job from a friend. We are ambivalent about the timing of Facebook friendship as compared to friend's start date because the time of Facebook friendship is only observable from 2007 onward, whereas the start date can take values before 2007. The figure shows the time period contact-based tie strength is measured over, as well as the point in time network overlap (or mutual friends) is measured.

network as given, we still encounter identification issues because tie strength is endogenously chosen.[25] Our identification issues are closely related to those

on the effect of the network as a whole, rather than the effect of tie strength between individuals in the network (Sacerdote, 2011; Beaman, 2012). These works have shown that the network alone, ignoring tie strength, has an effect on outcomes. We believe that the effect of exogenous networks and also exogenous tie strength is an important next step in this line of research.

[25]Tie strength is endogenous in a few ways. First, individuals may select into peer groups and tie strength in unobservable ways. Second, an individual may increase tie strength with a friend in the hopes of obtaining a job at that friend's employer. We are most concerned with the first issue, because we believe that we can control for strategic tie strength. Although contact based measures may be manipulated we believe that mutual friends is more difficult for an individual to fully control. For example, if Laura wants to

in the peer effects literature (Sacerdote, 2011; Manski, 1993).[26] The first candidate problem is "reflection" which occurs if an individual's outcome is a function of the average outcome of her network connections, and vice-versa. However, job transmission does not suffer from reflection because user i is transmitted her job from friend k, only if k has worked at their shared employer for at least one year. Thus, it is extremely unlikely for user i to transmit a job to friend k, and simultaneously for user k to transmit a job to user i.[27] Second, "correlated effects" occur when individuals select into peer groups in a way which is unobservable. Third, "contextual effects" point to the issue that background characteristics of the peer group may affect the outcome variable for both the individual and the peer. We will attempt to identify the relationship between tie strength and job transmission while controlling for correlated and contextual effects using individual level and dyad level controls, as well as a number of robustness checks.

Even if we are unable to fully control for these effects we can still say whether an individual is differentially likely to have a job transmission from a strong tie rather than a weak tie. But, this statement is made regardless of the reasons for that endogenously determined network of friendships and level of tie strength. To find the causal impact of tie strength on job transmission, one would ideally observe an exogenous shock to the network and tie strength, such

---

have more mutual friends with Jason, she can attempt to friend Jason's friends, but they may not accept her friendship. Also, she needs to encourage Jason to become friends with her friends, and furthermore convince her friends to accept Jason's friendship. Tags or posts from the user to a friend may be more easily manipulated by the user. However tags the other direction, from a friend to a user, are difficult to manipulate. In the Robustness Checks section we replicate our analysis using tags from the friend to the user. Also, one may believe that a user would only consider building strategic tie strength closer to the time of job search. We also replicate our results with differing time windows of measurement in the Robustness Checks section.

[26]Because our paper is not about the effect of peer group outcomes on individual outcomes, there is not a one-to-one mapping for the terms used in the peer effects literature to our problem. Thus, we may use the terms "reflection", "correlated effects" and "contextual effects" in a non-traditional manner.

[27]It is possible for user i's job transmission to affect future job transmissions to other friends j, l, and so on. But, because we take a random sub-sample at a given point in time this will not be recorded in our cross sectional data.

as an experiment that randomly assigns networks and tie strength. Keeping in mind the endogenous nature of our tie strength variables, this section begins by testing the Descriptive Weak Ties Hypothesis which mirrors the analysis from the seminal paper (Granovetter, 1973), and then proceeds beyond descriptive analysis to test the Conditional Weak Ties Hypothesis.

1.4.1 Descriptive Weak Ties Hypothesis

This section is purely descriptive in nature, so as to replicate the previous descriptive results. The analysis begins with tags because it is our strictest measure of contact-based tie strength; fewer dyads have tags than posts on Facebook.

Previous work has tested the Descriptive Weak Ties Hypothesis by identifying dyads in which a friend helped a person find a job and then showing the distribution of tie strength for those dyads. Generally, these papers found the distribution of tie strength has a large mass of weak ties and a smaller mass of strong ties. For example, Granovetter (1973) found that only 26% of his sample obtained their jobs through the help of a strong tie, a person they self-reported as seeing "often." Figure 1.5 shows the distribution of weak to strong ties using number of tags in a year as the tie strength measure for the 770,000 dyads for whom a job transmission occurred (X).[28] There is a huge mass of dyads who are weakly tied to each other. This same information is shown in log-log transformation in the bottom panels of Figure 1.5. The average number of tags between dyads with a job transmission is 0.8 tags per year, so over 91% of the dyads were transmitted a job from a friend with below average tags. Although most jobs come from weak ties, this does not necessarily mean an increase in the number of weak ties would cause more job transmissions. The distribution of tie strength is endogenously determined so we cannot make any causal inference about the effect of a change in this

[28]Recall the tags are measured for the full year before the user's most recent start date.

distribution on the level of job transmission.

Also pictured in Figure 1.5 are the distributions of tie strength for two alternative comparison groups. The previous paragraph discussed the tie strength distribution for the users and their friends who transmitted them a job.[29] Two logical comparison groups for these dyads come to mind. The first is those same users who were transmitted jobs and all their 30 million friends (Job Users & Friends, +). The second widens the sample to all the 296 million users and friends with available data (All, circle). All four panels of Figure 1.5 show the distribution of weak to strong ties for the transmission dyads only (X), and the left hand side compares this distribution to the same information for the job transmission users and all their friends (+), while the right hand side compares this same information to the distribution for all dyads (circle).

What is most striking about all these tie strength distributions is their similarity. Most jobs come from a weak tie, but using either comparison group, most ties in the population are weak. This means that the strength of weak ties is highly mechanical.[30] Although Figure 1.5 provides evidence in support of the Descriptive Weak Ties Hypothesis, it may not be that weak ties convey novel information. Rather, it could be that weak ties are important because most ties are weak. Although the distributions in Figure 1.5 are statistically significantly different from each other using a Wilcoxon signed-rank test, it is clear from a visual inspection that all distributions are characterized by many more weak than strong ties.[31] We summarize these findings below.

[29]These represent 400,000 users and their 770,000 friends.

[30]This type of tie distribution is not unique to Facebook nor to our measure of tie strength. Networks are often characterized by many more weak ties than strong ties, whether measured by mobile phone usage (Onnela et al., 2007), academic co-authorship (van der Leij and Goyal, 2011), Twitter usage (Grabowicz et al., 2012; Harrigan et al., 2012), or Facebook usage (Ferrara et al., 2012; Bakshy et al., 2012).

[31]For "All" vs. "Transmission Dyads Only" including 0 tag dyads: $z = -3.172$ $Prob > |z| = 0.0015$ and for "Job Users & Friends" vs. "Transmission Dyads Only" including 0 tag dyads: $z = -8.003$ $Prob > |z| = 0.0000$ using a Wilcoxon signed-rank test. When we restrict ourselves to the partial distribution that excludes dyads with zero tags (Figure 1.7 in the Appendix) we cannot reject the null hypothesis that the distribution of weak to strong ties is identical for job transmission dyads as compared to all dyads. For "All" vs. "Transmission Dyads Only" excluding 0 tag dyads $z = 0.947$ $Prob > |z| = 0.3435$ using

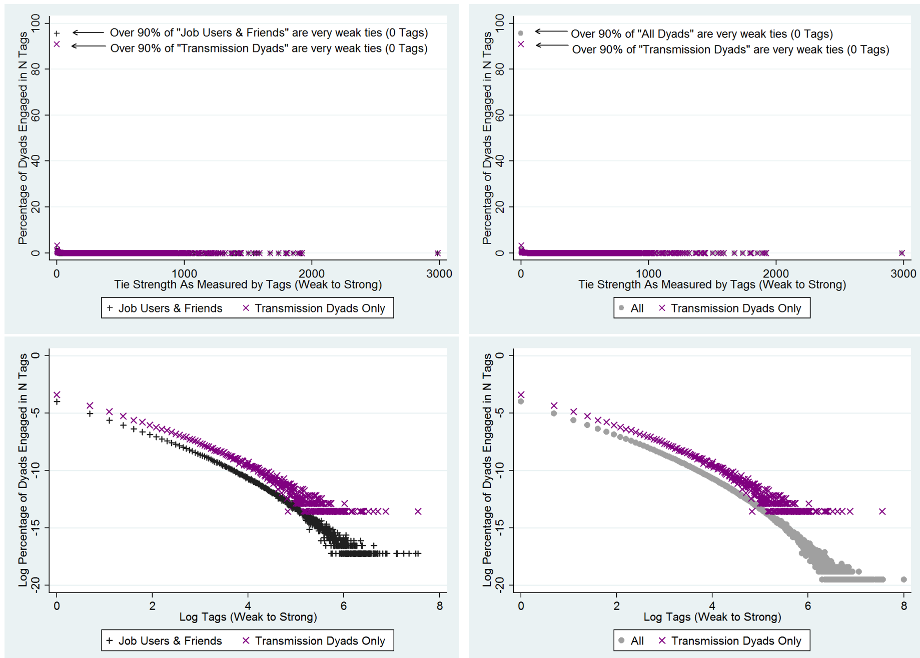


Figure 1.5: Distribution of Tags (including 0)

Note: The panels of this figure picture the distribution of weak to strong ties using number of tags from a user to a friend the year before the user began her most recent job as the measure of tie strength. The upper left panel shows the distribution for only those 770,000 dyads who had a job transmission (Transmission Dyads Only, X) versus the 30 million dyads where the user of interest was transmitted her job from a friend (Job Users & Friends, +). The upper right panel shows the same distribution for only those 770,000 dyads who had a job transmission (X) versus all the 296 million dyads with full information available in our data (All, circle). There are 400,000 users and 770,000 dyads in the "Transmission Dyads Only" (X) distribution. There are those same 400,000 users connected to all their 30 million friends in the "Job Users & Friends" (+) distribution. There are 6 million users and 296 million dyads in the "All" (circle) distribution. The lower panels show the same information in log-log transformation.

We conclude there is support for the Descriptive Weak Ties Hypothesis that most jobs will be found through a weak tie rather than a strong tie when this analysis is not conditioned on information about the network as a whole. In our data, over 90% of individuals are transmitted their job through a weak tie. However, this is largely driven by the fact that most ties are weak. A major strength of weak ties is that individuals have so many weak ties.

1.4.2 Conditional Weak Ties Hypothesis

This paper has shown that most jobs are transmitted by a weak tie, however most individuals have many weak ties. Thus far the analysis has been purely descriptive to more closely match the seminal paper (Granovetter, 1973). The original theoretical motivation for the importance of weak ties is they act as bridges that convey novel information, and not only their sheer numbers. To address this, we test whether tie strength negatively or positively predicts job transmission through a specific friend, conditional on information about the individual's whole network.

Conditional Weak Ties Hypothesis: Sub-sample

Job transmission J_{ik} is defined as a dummy variable that takes the value $J_{ik} = 1$ if a dyad meets our job transmission criteria, and takes the value $J_{ik} = 0$ otherwise.[32] A user i with N friends has information for each friend: $J_{i,k=1}$, $J_{i,k=2}$ to $J_{i,k=N}$. Job transmission $J_{ik} = 1$ is a rare occurrence between dyads. Conditional on the user of interest i having at least one occurrence of $J_{ik} = 1$, the average level of job transmission is about 2% between these users and all their friends.[33] To use information about a large set of users while still

a Wilcoxon signed-rank test For "Job Users & Friends" vs. "Transmission Dyads Only" excluding 0 tag dyads $z = 6.202$ $Prob > |z| = 0.0000$ using a Wilcoxon signed-rank test.

[32] Recall the job transmission criteria are: (a) user i and friend k have same employer, (b) user i starts work at least 365 days after friend k (c) user i and friend k became Facebook friends at least 365 days prior to start date.

[33] A dyad-level job transmission rate of 2% means that if a user i is transmitted her most recent job from a friend, on average 2% of her friends meet the job transmission criteria. For

maintaining a computationally manageable number of dyad-level observations, we took a random sample of 65,000 users from the 400,000 users who were transmitted a job from a friend. We then selected the friends of these 65,000 users where $J_{ik} = 1$ with 100% probability. The friends of these same 65,000 users where $J_{ik} = 0$ were sampled with only 10% probability.[34] The resulting one million dyads are our primary sub-sample. As a robustness check, the same analysis has been performed on an alternative sub-sample of a randomly selected 12,000 users and *all* of their friends. The results are very similar (see the Appendix for details). The primary sub-sample data is weighted so that each user's weights sum to one, ensuring individuals with many friends are treated similarly to individuals with very few Facebook friends.[35]

Table 1.1 presents the user-level summary statistics, and Table 1.2 contains the friend-level summary statistics which are quite similar. Both users and their friends are in their mid-20s, and the sample has slightly more women than men. The sample is well educated, with the highest level of education listed on their Facebook page as high school for 9%, college for 72%, and graduate school for 19%.[36]

---

example if she had 100 friends, then on average 2 friends would have transmitted her job to her, while if she had 200 friends, then on average 4 of her friends would have transmitted her job to her.

[34]This method has previously been used in the study of rare events (King and Zeng, 2001). See Manski and Lerman (1977) for related work on choice based sampling.

[35]See the Appendix Table 1.10 for an example of how the primary sub-sample was weighted. In the primary sub-sample most of the friends which are indexed by k show up only a single time in the data, so in text we report regressions without a friend fixed effect. We have also replicated the analysis with a friend fixed effect, and only using dyads where the friend shows up a single time. The results are very similar to those reported in text.

[36]Facebook users in our primary sub-sample are those who experienced a job transmission. One may wonder how they compare to all Facebook users with employer/education information, all Facebook users, and to the US population as a whole. The users in the primary sub-sample are slightly younger than all Facebook users with employer/education information, who in turn are about three-fourths the age of all Facebook users. In general Facebook users are younger than the average age of 37.5 reported by the CPS (Current Population Survey September 2012 http://www.census.gov/cps/). The primary sub-sample has slightly more women than men, which is also true of the Facebook users with employer information, all US Facebook users, and data from the CPS. The Facebook users in our sample all have at least some high school. The proportion of US Facebook users that list

Table 1.1: User Summary Statistics (Primary Sub-sample)

Variable	Mean	Std. Dev.	Min.	Max.
age	24.598	5.402	16	65
male	0.474	0.499	0	1
married	0.188	0.391	0	1
single	0.29	0.454	0	1
in relationship	0.256	0.436	0	1
engaged	0.061	0.239	0	1
some HS	0.09	0.286	0	1
some college	0.72	0.449	0	1
some post BA	0.19	0.393	0	1
friend count	524.032	341.899	0	5054

N=1,017,089 Dyads
65,590 Users
Each dyad-level observation is weighted
by 1/(number of times user is in data)

Table 1.2: Friend-level Summary Statistics (Primary Sub-sample)

Variable	Mean	Std. Dev.	Min.	Max.
F age	25.232	5.74	16	65
F male	0.476	0.499	0	1
F married	0.211	0.408	0	1
F single	0.272	0.445	0	1
F in relationship	0.249	0.433	0	1
F engaged	0.059	0.237	0	1
F some HS	0.088	0.284	0	1
F some college	0.723	0.448	0	1
F some post BA	0.189	0.392	0	1
F friend count	503.027	389.244	0	6220

N=1,017,089 Dyads
861,380 Friends
Each dyad-level observation is weighted
by 1/(number of times friend is in data)

some high school is similar to the 21% of respondents to the CPS who report having some high school. Last, the CPS finds that 41% of respondents are married while the proportion of individuals on Facebook that elect to list themselves as married is below that found in

Conditional Weak Ties Hypothesis: Structure-based and Contact-based Tie Strength Measures

The goal is to find the relationship between the tie strength measures and job transmission conditional on information about a person's whole social network. Table 1.3 summarizes the statistics for job transmission and tie strength in the primary sub-sample. The average level of job transmission J_{ik} in our random sub-sample of one million friendships is 1.8%. Our contact-based measures of tie strength are tags and posts. A very small number of dyads, only 4.3%, have any tags between them, while many more dyads, 16%, have some posts. On average, a dyad has 54 mutual friends, almost all dyads have some network overlap, and on average a dyad shares about 5.2% of their friends as measured by network overlap.

For the scaled measures of tags and posts, we prefer the alternative sub-sample with 12,000 users and all of their friends. The measures for percentage of tags or posts are estimated in the primary sub-sample whereas they are the true percentages in the alternative sub-sample.[37] For our alternative sub-sample, the average level of percentage of tags (including zeros) is 0.6% and when we exclude dyads with zero tags it is 14.7%. The average level of percentage of posts (including zeros) is 0.8% and rises to 5.1% when excluding dyads with zero posts. All the summary statistics for the alternative sub-sample are in the Appendix in Tables 1.17, 1.18, 1.19, and 1.21.

Table 1.3 does not elucidate the relationship between contact-based and structure-based tie strength measures. Recall that there is evidence of a diversity-bandwidth tradeoff in which dyads with more novel information, as measured by low levels of structure-based tie strength, may also have lower

the CPS.

[37]In the primary sub-sample we only observe tags and posts for 10% of the $J_{ik} = 0$ dyads, so we cannot observe the true total number of tags or posts for a user i. To estimate the total tags and posts for a user i we sum the contact for dyads with $J_{ik} = 0$ and multiply by $\frac{1}{10\%}$ and add this to the sum of the contact for dyads with $J_{ik} = 1$. We use this total as our estimate of the denominator for percentage of tags and posts reported for the primary sub-sample.

Table 1.3: Dyad-level Tie Strength Summary Statistics (Primary Sub-sample)

Variable	**Mean**	**Std. Dev.**	**Min.**	**Max.**	**N**
job transmitted	0.018	0.132	0	1	1,017,089
any dyad tag	0.043	0.203	0	1	1,017,089
tags	0.297	3.951	0	816	1,017,089
tags (1+)	6.859	17.77	1	816	50,126
% of tags*	0.003	0.027	0	1	1,017,089
% of tags (1+)*	0.069	0.111	0	1	50,126
any dyad post	0.163	0.369	0	1	1,017,089
posts	0.374	2.283	0	619	1,017,089
posts (1+)	2.299	5.253	1	619	178,705
% of posts*	0.006	0.029	0	1	1,017,089
% of posts (1+)*	0.036	0.063	0	1	178,705
mutual friends	54.298	59.154	0	2074	1,017,089
any network overlap	0.985	0.121	0	1	1,017,089
network overlap	0.052	0.052	0	1	1,017,089

1,017,089 Dyads
Weighted by inverse of sampling probability
* Estimated scaled measures for the primary sub-sample

bandwidth, as measured by contact-based tie strength. Table 1.4 shows the positive but weak correlation between structure-based and contact-based tie strength. Table 1.5 reports the correlation between the scaled measures for the alternative sub-sample.[38] These tie strength variables move in the same direction, but neither tags nor posts is a great proxy for mutual friends or network overlap.

Table 1.4: Correlation Absolute Tie Strength (Primary Sub-sample)

Variables	mutual friends	tags	posts
mutual friends	1.000		
total dyad tags	0.034	1.000	
total dyad posts	0.036	0.332	1.000

[38]The correlation is similar for the primary sub-sample which is presented in Table 1.25 of the Appendix.

Table 1.5: Correlation Percentage Tie Strength (Alternative Sub-sample)

Variables	network overlap	pct. tags	pct. posts
network overlap	1.000		
percentage tags	0.057	1.000	
percentage posts	0.053	0.267	1.000

Conditional Weak Ties Hypothesis: Empirical Specification

To test the Conditional Weak Ties Hypothesis, we estimate the relationship between job transmission J_{ik} and tie strength. As aforementioned, we are concerned with contextual and correlated effects when estimating this relationship. Individuals and their friends may have correlated observable and unobservable variables which affect both tie strength and job transmission. To control for these effects we first include a user fixed effect. The user (ego) fixed effect, E_i, controls for all observable and unobservable attributes about the individual. For example, an extroverted individual may be more likely to have a job transmission, and to have higher levels of tie strength. With the inclusion of E_i, the variation in tie strength, T_{ik}, comes from variations within a user's friendships instead of across all dyads.

Additionally, dyad-level variables may also affect tie strength and job transmission. It is likely that dyads with higher tie strength may be more likely to be very similar in other unobservable ways, homophily. Because of this homophily, these dyads may be more likely to work at the same employers, even in the absence of job transmission through social networks. Ideally, we would use a dyad fixed effect to control for this source of omitted variable bias, but we only observe each dyad a single time. The Facebook data, however, have a rich set of dyad-level control variables summarized in Table 1.6 that we include in our analysis.[39] Previous work has shown that gender, age, race and

[39]We were also able to identify the distance in kilometers and the industry for some of the dyads in our sample. Including these additional controls for homophily does not change the sign or significance of our results. Please see the Robustness Checks section for this

education differences are all predictive for labor market outcomes or decisions in general (Aral and Walker, 2012; Lin, 1999; Leicht and Marx, 1997; Holzer, 1987). We are able to control for all of these with the exception of race, as well as the following other observable dyad variables: friend's tenure at employer at time of user's most recent start date, same relationship status, and same city/state.[40]

Also, it is possible that friend specific unobservable attributes affect job transmission. For example, a friend may have the unofficial position of recruiter for her firm. This would increase the likelihood of job transmission from that friend, but is unobservable. Because we use a random sample of dyads, most friends only occur one time in the data so we cannot use a friend fixed effect to control for friend level unobservables.[41] We include a user fixed effect and dyad-level controls which are computed from differences between the user and friend-level variables, so we cannot include many friend-level variables in our analysis. For example, including both friend's age, and age difference is akin to including the same variable twice in our model. We do include the number of friends of the alter raised to the fourth power to flexibly control for the the friends of a friend.[42]

We are left with the following linear model:

$$J_{ik} = \beta T_{ik} + \alpha X_{ik} + \Gamma A_k + E_i + \epsilon_{ik}$$

A user i with N friends has a job transmission dummy for each of those

---

analysis.

[40]Users in our data list their relationship status as married, engaged, single, or in a relationship. The variable same relationship status takes the value one if a user and friend both have the same status listed. We have also used specifications with this information coded as both married, both single etc. The variable same state (same city) takes the value one if both users reside in the same US state (city) during the year of the user's most recent start data.

[41]We have replicated the same analysis with a friend fixed effect for those friends who occur two or more times in the data. The results are extremely similar, so we do not include them in text.

[42]We tried a number of different functional forms to control for the number of alter's friends and all had very similar results. So we only include the functional form with alter's friends raised to the 4th power in text.

friends: $J_{i,k=1}$, $J_{i,k=2}$ up to $J_{i,k=N}$. The dummy variable J_{ik} takes the value 1 if person i was transmitted her most recent job from friend k, T_{ik} is a vector of tie strength variables, X_{ik} is a vector of dyad-level control variables, A_k is a vector of the number of alter's friends raised to the fourth power, and E_i is the user (ego) fixed effect. The standard errors are clustered at the user i level.[43] The dependent variable takes the values zero or one, so a conditional logit model would be appropriate. However, we are most interested in the average probability of job transmission, and the coefficients from the linear model are more easily interpreted. So, we ignore the special nature of the dependent variable and report the results from the linear model in text.[44]

In this specification, β is the average percentage point difference in the likelihood of job transmission attributable to a unit increase in T_{ik}, tie strength. The use of a user fixed effect allows our model to control for the idea that an extrovert may be more likely to have high levels of tie strength,

[43]Clearly ϵ_{ik} is not independent and identically distributed because we have multiple observations within each individual i which are not independent of each other. To correct for this we have clustered our standard errors at the individual-level in addition to our inclusion of a user fixed effect. There is the additional worry that ϵ_{ik} for individual i may be correlated with ϵ_{jh}, however the use of a random sub-sample of only 1 million dyads mitigates this concern. The Facebook network is highly connected in general, so it is very possible for a random person i and another random person j to interact with each other. However, because we are only looking at a random selection of 65,000 individuals out of 400,000 who had a job transmission (who in turn are from a total of 6 million individuals) we believe it is reasonably safe to assume that in our sub-sample ϵ_{ik} is uncorrelated with ϵ_{jh}. The cluster-robust standard error estimator converges to the true standard error as the number of clusters (65,000 individuals in our data) not the number of observations (1 million dyads) approaches infinity. This is another advantage of our sampling procedure because we have been able to sample more individuals (number of clusters). It is likely that 65,000 clusters of roughly equal size is large enough for accurate inference (Kezdi, 2004).

[44]Another reason we present the linear model in the body of the paper is that our primary sub-sample data has weights which vary within the individual. Recall we have sampled an individual's friends with $J_{ik} = 1$ with 100% probability and their friends with $J_{ik} = 0$ with only 10% probability. We have included a user-level fixed effect in our analysis, but a conditional logit model calls for the weights to be equal within the level of the fixed effect. So we cannot actually use a conditional logit model on our primary sub-sample. We do however present the results for our alternative sub-sample of 12,000 users and all their friends in the Appendix. The results from the linear fixed effects model across both sub-samples are extremely similar. Additionally, the results from the fixed effects linear model versus the conditional logit model for the alternative sub-sample are also extremely similar. All these results are reported in the Appendix.

Table 1.6: Dyad-level Demographic Summary Statistics

Variable	Mean	Std. Dev.	Min.	Max.	N
F years older (10)	0.047	0.454	-4.5	4.9	1,017,089
both male	0.25	0.433	0	1	1,017,089
both female	0.292	0.454	0	1	1,017,089
F more educated	0.156	0.363	0	1	1,017,089
F less educated	0.194	0.395	0	1	1,017,089
F more friends (100)	-1.29	5.311	-49.42	54.88	1,017,089
same relationship status	0.284	0.451	0	1	1,017,089
same state at start date	0.475	0.499	0	1	1,017,089
same city at start date	0.177	0.382	0	1	1,017,089
F tenure at employer (years)	1.127	1.778	-3.751	43.112	1,017,089
same high school	0.294	0.455	0	1	1,017,089
same college	0.303	0.459	0	1	1,017,089
same grad school	0.016	0.126	0	1	1,017,089
1,017,089 Dyads					
Weighted by inverse of sampling probability					

and more likely to have a job transmission. The use of dyad-level controls allows our model to control for the idea that a user and friend who are both 16 year old may have higher tie strength, may both be more likely to work at McDonalds, and may be more likely to have a job transmission. However, we may still not be identifying the causal relationship between tie strength and job transmission, because T_{ik} is not exogenously determined.

Before presenting the results from our basic specification, we would like to reemphasize that our results may not be the true *causal* effects of tie strength on job transmission. The principal concern in using our estimating equation is that tie strength T_{ik} is endogenous, $E[T_{ik}\epsilon_{ik}] \neq 0$. Even with this endogeneity, we can confidently make statements about the correlation between tie strength and job transmission above unobservable individual heterogeneity and observable dyad heterogeneity. However this correlation may not be the true causal β which would be the impact of exogenously increasing tie strength on job transmission.

For example, a dyad with a friend who is an extrovert might be more likely to have a job transmission, as well as a higher number of of mutual friends, tags, and posts. Recall, we have controlled for the user being an extrovert by including a fixed effect, but we are unable to control for the alter being an extrovert. If having a friend who is an extrovert, regardless of the user being an extrovert, increases tie strength and job transmission, then one would expect the coefficient on tie strength, β, to be biased upward from the true causal β. Another scenario is that unproductive individuals may be less likely to have a job transmission but may be more likely to have high tie strength. Again, we have controlled for the user being unproductive with a fixed effect. However, if having an unproductive friend, regardless of the user being unproductive, increases tie strength but decreases job transmission, we would expect our estimates of β to biased downward from the true causal β. If ϵ_{ik} is catching these types of confounding factors, our estimates of β may be biased away from the true causal effect of tie strength, and it is unclear the direction of this bias.

The most compelling way to control for endogeneity would be to experimentally assign tie strength to dyads. Secondarily, we could use an exogenous shock to tie strength in an instrumental variables approach. We did find a possible instrument for number of mutual friends in our data in the form of a randomized experiment run by Facebook. We need an instrument to be both valid and relevant, so as a first step we tested for relevance. That is we tested if $E[T_{ik}Z_{ik}] \neq 0$ where T_{ik} is our measure of tie strength and Z_{ik} represents our instrumental variable. Unfortunately the relationship between random assignment into the experimental treatment and tie strength was very weak, so we did not pursue this analysis any further.

Conditional Weak Ties Hypothesis: Results

In Table 1.7 we report the coefficients from our linear model where we look at each tie strength measure alone.[45] In columns 1 to 3, there is a dummy variable for if the dyad had any tags, posts or mutual friends. Think of these

[45] See Table 1.22 in the Appendix for the full results of Table 1.7 in text.

as measures of movement along the extensive margin of tie strength, meaning going from no measurable tie strength to any positive amount of tie strength. The coefficients in columns 1 to 3 are all positive and statistically significant. A movement from no tags to any tags is correlated with a 2.5 percentage point increase in the likelihood of job transmission with that particular friend. A movement from no posts to any posts is correlated with a 1.7 percentage point increase in the likelihood of job transmission. Last, a movement from no mutual friends to any mutual friends is correlated with a 0.8 percentage point increase in the likelihood of job transmission. Although these numbers may seem small, this means that a dyad with any tags has a probability of job transmission about 2.5 percentage points higher than a dyad with no tags which is a very large increase from the baseline average transmission rate of 1.8%.

Table 1.7: Linear Models Dependent Variable Job Transmission Absolute Tie Strength Measures One-by-One (Primary Sub-sample)

Variable	Exten. Tag 1	Exten. Post 2	Exten. M Frnd 3	Inten. Tag 4	Inten. Post 5	Inten. M Frnd 6	Ext/Int Tag 7	Ext/Int Post 8	Ext/Int M Frnd 9
any dyad tag	0.025***						0.022***		
	(0.001)						(0.001)		
any dyad post		0.017***						0.014***	
		(0.000)						(0.001)	
any mutual friends			0.008***						0.007***
			(0.002)						(0.002)
tags (10)				0.008***			0.004***		
				(0.001)			(0.001)		
posts (10)					0.019***			0.012***	
					(0.002)			(0.002)	
mutual friends (100)						0.007***			0.007***
						(0.000)			(0.000)
R2	0.121	0.122	0.120	0.121	0.121	0.121	0.121	0.122	0.121
N	1,017,089	1,017,089	1,017,089	1,017,089	1,017,089	1,017,089	1,017,089	1,017,089	1,017,089

Includes user-level fixed effect. Standard errors clustered at the user-level. Weighted so that each user's weights sum to 1. Control variables: friend x years older, both male/female, friend more/less educated, same relationship status, same city/state, friend's tenure at employer (years), same high school/college/graduate school, alter's number of friends, alter friends^2, alter friends^3 and alter friends^4. Legend: * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$

The next three columns of Table 1.7 look at continuous measures of tie strength which include all the zero tags, zero posts and zero mutual friends dyads. In columns 4 through 6, all the coefficients are positive and significant. When we decompose individual tie strength measures into the extensive and intensive margin we see that it both matters if any tie strength exists (extensive) and the level of that tie strength (intensive). We read the results in column 7 of Table 1.7 as saying that compared to an identical dyad with zero tags, a dyad with one tag has a 2.2 percentage point higher probability of job transmission, and a dyad with 10 tags has an additional 0.4 percentage points higher probability of job transmission. A dyad with 10 tags has a probability of job transmission about 2.6 percentage points higher than an identical dyad with no tags. The same pattern is illustrated for posts in column 8, and mutual friends in column 9.

Table 1.8: Linear Models Dependent Variable Job Transmission Scaled Tie Strength Measures One-by-One (Alternative Sub-sample)

Variable	**Exten. Tag 1**	**Exten. Post 2**	**Exten. M Frnd 3**	**Inten. Tag 4**	**Inten. Post 5**	**Inten. M Frnd 6**
any dyad tag				0.011***		
				(0.002)		
any dyad post					0.009***	
					(0.001)	
any mutual friends						0.009***
						(0.003)
% of tags	0.094***			0.074***		
	(0.008)			(0.010)		
% of posts		0.124***			0.093***	
		(0.011)			(0.013)	
network overlap			0.095***			0.093***
			(0.007)			(0.007)
R2	0.132	0.132	0.131	0.132	0.132	0.131
N	1,438,699	1,438,699	1,438,699	1,438,699	1,438,699	1,438,699

Includes user-level fixed effect. Standard errors clustered at the user-level.
Weighted so that each user's weights sum to 1.
Controls: F years older, both male/female, F more/less educated, same relationship status, same city/state, F tenure at employer (years), same high school/college/graduate school alter friends, alter friends^2, alter friends^3 and alter friends^4
Legend: * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$

Table 1.8 reports the results from linear models with the scaled measures of tie strength from our alternative sub-sample.[46] This is our preferred sub-sample for analysis of the scaled measures of tags and posts because we do not have to estimate the scaled measures in the alternative sub-sample. We interpret column 1 of Table 1.8 as stating that a dyad with 10 percent more tags will have a 0.9 percentage point higher probability of job transmission than a dyad that is identical in all other ways.[47] In columns 2 and 3 of of Table 1.8, we see that higher levels of scaled posts or network overlap are also associated with a higher probability of job transmission. Columns 4 to 6 of Table 1.8 show this same pattern holds if we decompose our scaled measures into the extensive and intensive margin. The results using the primary sub-sample are the same sign and significance for columns 1, 2, 3, and 6.[48]

All of these models are evidence against the Conditional Weak Ties Hypothesis. When we condition our analysis on information about an individual and all her friends we find a statistically significant and positive relationship between all our tie strength measures and job transmission. This is surprising because it is exactly the opposite of the implication from the original intuition that weak ties are best for conveying useful job information.

Although all three of our tie strength measures are positively correlated, there is reason to believe that each contributes something distinct to our understanding of the relationship between tie strength and job transmission. Tags may measure real world contact, while posts measure online contact, and

[46]The full regression results for the models in Table 1.8 are in Table 1.23 of the Appendix.

[47]A dyad with 100 percent more tags (e.g. a dyad with a user who only tags a single friend vs. a dyad without any tags) would have a 9.4 percentage points higher probability of job transmission.

[48]In Table 1.24 of the Appendix we report the results of models including our estimated scaled measures of tie strength from the primary sub-sample and show that sign and significance on these scaled measures in columns 1, 2, 3, and 6 are all positive and significant. Although, in columns 4 and 5 the coefficients on the "any dyad tag" and "any dyad post" are negative, the size of these coefficients (-0.068 and -0.038) are much smaller than the coefficients on the intensive margin variables "% tags" and "% posts" (1.009 and 0.990). We believe that noise in the estimated measures of percentage tags and percentage posts is causing the difference in results across the two samples.

mutual friends measure the structure of the network. Tags and posts measure contact in the past year and may be more easily manipulated by the individual, whereas the number of mutual friends is a long term level of closeness which is difficult for the user to fully control. To test if all three measures account for the same underlying closeness, we put all three into the same model. In Table 1.9 we interpret the coefficients on each tie strength variable as the contribution of that particular tie strength measure holding the other measures constant.[49] We interpret column 1 of Table 1.9 as saying compared to a dyad with no tags, no posts, and no mutual friends an identical dyad with some tags has a 1.9 percentage point higher probability of job transmission (holding constant posts and mutual friends); an identical dyad with some posts has a 1.4 percentage point higher probability of job transmission (holding constant tags and mutual friends); an identical dyad with some mutual friends has a 0.7 percentage point higher probability of job transmission (holding constant tags and posts). Columns 2 and 3 of Table 1.9 also indicate a positive relationship between the continuous absolute tie strength measures and job transmission. In column 4 and 5 of Table 1.9, we present the coefficients on the scaled measures of tie strength from our alternative sub-sample.[50] We see that these also have a positive relationship with job transmission.

We conclude there is strong evidence against the Conditional Weak Ties Hypothesis which states that the probability that an individual obtains a job through a particular social tie decreases with tie strength. Increases in contact-based tie strength (tags, posts) or network structure-based tie strength (mutual friends) are associated with a higher probability of job transmission from that

[49]The full results for Table 1.9 are in Table 1.26 in the Appendix for the full results. The same models replicated with the other data set are available in Table 1.15 in the Appendix.

[50]For the results from the primary sub-sample with the estimated scaled measures see columns 4 and 5 of Table 1.15 in the Appendix. In the model in column 4 using the primary sub-sample the coefficient on network overlap is insignificant, and in column 5 the coefficients on any tag or any post are negative. We believe this is driven by the fact that we are only estimating these scaled measures in the primary sub-sample. The result for the models of the form in columns 1 to 3 for the alternative sub-sample are also in Table 1.15 in the Appendix.

Table 1.9: Linear Models Dependent Variable Job Transmission All Tie Strength Measures

Variable	Exten. Absolute 1 Primary Sample	Exten. Absolute 2 Primary Sample	Ext./Int. Absolute 3 Primary Sample	Inten. Pct. 4 Alt. Sample	Ext./Int. Pct. 5 Alt. Sample
any dyad tag	0.019***		0.015***		0.005**
	(0.001)		(0.001)		(0.002)
any dyad post	0.014***		0.012***		0.007***
	(0.000)		(0.001)		(0.001)
any mutual friends	0.007***		0.006***		0.009**
	(0.002)		(0.002)		(0.003)
tagging (10)		0.005***	0.002***		
		(0.001)	(0.001)		
posts (10)		0.016***	0.008***		
		(0.002)	(0.001)		
mutual friends (100)		0.006***	0.005***		
		(0.000)	(0.000)		
% of tags				0.069***	0.058***
				(0.008)	(0.010)
% of posts				0.098***	0.072***
				(0.011)	(0.013)
network overlap				0.075***	0.068***
				(0.007)	(0.007)
R2	0.122	0.121	0.122	0.133	0.133
N	1,017,089	1,017,089	1,017,089	1,438,699	1,438,699

Includes user-level fixed effect. Standard errors clustered at the user-level. Weighted so that each user's weights sum to 1. Control variables: F years older, both male/female, F more/less educated, same relationship status, same city/state, F tenure at employer (years), same high school/college/graduate school, alter friends, alter friends^2, alter friends^3 and alter friends^4 Legend: * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$

friend when we control for contextual and correlated effects using an individual-level fixed effect and dyad-level control variables.

1.4.3 Robustness Checks

The previously presented results hold under a number of robustness checks.

Robustness Checks: Sensitivity Analysis

We use a number of controls in testing the Conditional Weak Ties Hypothesis. We test for sensitivity to these control variables, by comparing models with no controls to models in which we add a set of controls one by one. Our first set of controls is demographic, the second set is educational, the third is friend's tenure, and the fourth is the alter's friends.[51] The results in Table 1.11 of the Appendix show a positive and significant coefficient on tags and posts in the models with no controls. The coefficients on tags and posts remain positive as we add controls, but drop very slightly. This suggests the omitted variable bias is biasing these coefficients slightly upwards.[52] We find a negative and significant coefficient on mutual friends without controls, but as we add controls the coefficient becomes larger and positive, suggesting our omitted variable bias may be large and biasing this coefficient downward.[53] This possible downward bias on the mutual friends coefficient only strengthens the robustness of the positive coefficient found using all the observable variables. We conclude the positive relationship between tie strength and job transmission is robust to this sensitivity analysis.

Robustness Checks: Non-linear Specifications

In testing our Conditional Weak Ties Hypothesis we used linear models. But, in order to explore more complex relationships in the data, we also divided each of the tie strength measures into roughly equally sized quantiles (excluding zero). We estimate the coefficients from a linear model including these bins of tie strength, a user fixed effect, and the control variables. By plotting the coefficients against the level of tie strength, we obtain a visual representation of the relationship between tie strength and the probability of job transmission without imposing linearity. In all cases, the relationship is positive and generally linear, as pictured in Figure 1.8 in the Appendix.[54]

[51]Demographic controls are: friend years older, both male/female, same relationship, and same city/state. Education controls are: friend more/less educated, same high school/college/graduate school. We find some interesting sensitivity to the inclusion of the same high school variable which relates to other research about tie age (McEvily et al., 2012; Baum et al., 2012).

[52]For tags the coefficient with no controls is 0.009 and with all controls only falls to .008. For posts the coefficient with no controls is 0.020 and only falls to 0.019 with all controls.

[53]The coefficient with no controls is -0.001 on mutual friends, and with all controls it rises to 0.007.

[54]The underlying regression results are quite long, so they are available from Laura K. Gee (l1gee@ucsd) by request.

Also, job transmission takes either the value zero or one. In the main text, we reported the results from linear models. However, given the special nature of the dependent variable, we also use a conditional logit model to estimate the relationship between tie strength and job transmission. We assume that the likelihood that a friend k transmits a job to a user i is determined by characteristics that vary by user and friend. We can write that relationship as follows:

$$V_{ik} = v_{ik} + \epsilon_{ik}$$

$$V_{ik} = c + \beta T_{ik} + \alpha X_{ik} + \Gamma A_k + \eta E_i + \epsilon_{ik}$$

Let c be a constant, T_{ik} be a vector of tie strength variables, X_{ik} be a vector of dyad control variables, A_k be a vector of alter control variables, E_i be a user-level fixed effect and ϵ_{ik} be the error term. Let F_i denote the set of friends that user i has. If we assume that the error terms follow an extreme value distribution, the probability that a user is transmitted a job from friend k can be expressed as:

$$P_{ik} = \frac{e^{v_{ik}}}{\sum_{m \in F_i} e^{v_{mk}}}$$

The conditional logit results are reported in Table 1.16 in the Appendix. The coefficients in all the conditional logit models are of the same sign and significance as those presented from the linear model, so our results are robust to non-linear specifications.[55]

[55]The coefficient in the logit models are difficult to interpret. In all our models the coefficient on "F years older" is negative, indicating that the probability of job transmission is negatively correlated with the number of years older a friend is than the user. To help understand the magnitude of the coefficients in the logit model we take the coefficient on the tie strength measure multiplied by its standard deviation and divide by the coefficient on age difference. Then we can say that for a one standard deviation increase in tags, a person would befriend a person .4 years older than them. For a one standard deviation increase in posts, a person would befriend a person .8 years older than them. Also, for a one standard deviation increase in mutual friends, a person would befriend a person 3.5 years older than them. The median age difference in the sub-sample is 0 years.

Robustness Checks: Incidental Job Transmission

The job transmission variable may include incidental occurrences of two individuals working at the same workplace. Because these incidental job transmissions may be especially likely for large employers, we replicate our analysis excluding large employers.[56] For the Descriptive Weak Ties Hypothesis, we initially found that 91% of our job transmissions came from a tie with zero tags. If we exclude employers who were listed by over 10,000 users in our 6 million user sample, we find 90% of the job transmissions come from dyads with zero tags. Even, if we restrict the sample to users with an employer with 10 or fewer mentions, we still find that 86% of job transmissions come from dyads with zero tags between them.[57] To address the same concern in testing the Conditional Weak Ties Hypothesis we exclude users who list an employer who is in the top 25% of employers by size in our primary sub-sample. Yet, we still find the same sign and significance on the tie strength coefficients in Table 1.12 in the Appendix. As an additional check, the sample is restricted to users with an employer with only a single mention in the data. Again there is no change in the sign or significance of the results.

Individuals who work for the same firm in different cities may be more likely to have an incidental job transmission, so we redefine our dependent variable J_{ik} to only take the value one if a user and a friend live in the same city. For the Descriptive Weak Ties Hypothesis, we find that with this more restrictive job transmission definition there is still skew toward weak ties with 88% of jobs being transmitted by a dyad with zero tags. Also for the Conditional Weak Ties Hypothesis, we still find a positive and significant relationship between tie strength and job transmission with the exception of the

[56]The largest employer in our primary sub-sample was listed by 39,690 users, which is only 0.6% of this sample. We initially found that 7% of our users experienced a job transmission. If we limit the sample to employers with 1,000 or fewer mentions that falls to 6%. If we further limit to employers with 100 or fewer mentions that falls to 5%.

[57]If we limit to employers with fewer than 1,000 users, the proportion from zero tag dyads falls to 89.8%. For only employers with 100 users or less, the proportion is 88.5%.

loss of significance on the any mutual friends dummy variable in Table 1.12 in the Appendix.[58] This more restrictive definition excludes two individuals who work for the same employer but live in neighboring cities. For example, two professors who work at UC San Diego with one living in La Jolla and the other living in San Diego. We identify distance between a user and a friend for 30% of the sample and add a dummy for distance identification as well as an interaction with distance into the model.[59] There is no change in the sign or significance of the coefficients on tie strength in Table 1.12 in the Appendix when we include distance in our original models.

We observe many job transmissions at employers that do not generally require a college education, and one may believe that these jobs are more prone to incidental job transmission. For the Descriptive Weak Ties Hypothesis we restrict the analysis to only users with some college, and find that 93% of these users are still transmitted a job through a tie with zero tags.[60] For the Conditional Weak Ties Hypothesis, we restrict the analysis to only users with some college, then only users with some graduate school, and last to only those dyads where both the user and the friend have some college. In column 6 of Table 1.12 in the Appendix, the model with only graduate school users, we find the coefficient on the mutual friends dummy loses significance. However, in all other cases, the sign and significance of the coefficients on tie strength remain the same as the models in text.[61]

Additionally, certain industries may be especially prone to incidental

[58]Keeping in mind that the sample for this model is only 369,802 dyads and 98% of the dyads have mutual friends we do not believe this loss of significance is an issue for the robustness of our findings.

[59]We are only able to identify distance for 30% of the sample for a few reasons. First, we can only identify the city name and state for each individual, but there are multiple US locations with the same city name within a state. For example there are three cities named Oakwood in Ohio. Second, we matched our city state pairs to a list of city state pairs, and this list did not always contain the distances between smaller cities.

[60]When we restrict the sample to all college users and friends 96% of dyads have zero tags, so it is still the case that most ties are weak in general.

[61]Keeping in mind that the sample for the model where mutual friends loses significance has only 213,724 dyads and 98% of the dyads have mutual friends we do not believe this loss of significance is an issue for the robustness of our findings.

job transmission, so we match the self-reported employer names from Facebook to industries.[62] We replicate the analysis while including a dummy for identifying industries and an interaction with a dummy variable for the same industry, and all tie strength coefficients remain positive and significant. Also, the coefficients on tie strength remain positive and significant if we include a dummy for identifying the friend's industry and dummy variables for the friend's industry from the 20 industry classifications in the NAICS.

Robustness Checks: Definition of Tie Existence and Tie Strength

In this section, we show that the results are robust to varying definitions of tie existence and tie strength. A major concern is that dyads without any contact are too weak to be considered friends. As a robustness check on both our hypotheses we redefine a tie as only existing if some contact has occurred. For the Descriptive Weak Tie Hypothesis we compute the distributions of weak to strong ties for only those dyads with a positive (1+) amount of tags between them as shown in Figure 1.7 of the Appendix.[63] Under this definition of the network, the distribution is still highly skewed toward weak ties even when we have removed those with no tags from the analysis. For the Conditional Weak Tie Hypothesis we limit the sample to dyads with at least one tag or post in the previous year, and all the coefficients on tie strength that remain

[62] We use the ReferenceUSA database to find the North American Industry Classification System (NAICS) industry for employers with 500 or more employees listed in the ReferenceUSA database. Because this matching involves linking to an outside data set we are only able to do this for the sub-sample used in the Conditional Weak Ties Hypothesis. See www.referenceusa.com and http://www.census.gov/epcd/www/sic.htmlandhttp://www.census.gov/eos/www/naics/. Because employer names are self-reported on Facebook and these data are quite inconsistent, we are only able to match both a user's and friend's industry for 14% of the primary sub-sample. We use the unique employer ID number assigned by Facebook to each self-reported employer name to compute job transmission, so the issue of self-reporting does not affect our job transmission variable in the same way that it affects matching to an outside data source. Using only the employer ID and not the written name to compute job transmission will understate the number of job transmissions in our data.

[63] We use only tags to mirror our earlier analysis.

significant also remain positive as detailed in Table 1.13 of the Appendix.[64]

An individual may self-select into higher contact-based tie strength with a friend in the hopes of obtaining a job from that friend.[65] If this is the case then we capture the effect of strategic tie strength, rather than underlying tie strength on job transmission. In choosing the time frame for measuring contact-based tie strength there is a tradeoff between how current the measure of tie strength is and how likely the tie strength is strategically motivated. In the text, tags and posts are measured from a user to a friend during the year previous to the user's most recent start date.[66] As a robustness check we have replicated the analysis for the Conditional Weak Ties Hypothesis using two additional time frames. First, we measure tags from a user to a friend for a year starting two years before the user's start date.[67] Second, we compute tags from a user to a friend for one year starting one month before the user's most recent start date.[68] Both these measures have the added benefit of excluding tags or posts that may be the consequence of knowing that a user will work with a friend. As an additional check on strategic tie strength we compute tags from a friend to a user, and bi-directional tags. Both of these measures are more difficult for a user to manipulate in the hopes of obtaining a job transmission. In all cases, the coefficients on the tag measures in Table 1.13 remain the same sign and significance.

It is interesting that we find such robust support for a positive relationship between tie strength and job transmission because there is a strong intuitive argument for the importance of novel information from weak ties. A

[64]The "any post" and "any mutual friend" dummies lose significance in this small sample of 131,175 dyads, most likely because 90% of these dyads have a post and 98% have some mutual friends.

[65]This is less likely for the network-structure based measure of mutual friends.

[66]If a person started a job in September 1, 2012, then tags are measured from September 1, 2011 to the day before September 1, 2012.

[67]If a person started a job in September 1, 2012, then tags are measured from September 1, 2010 to the day before September 1, 2011.

[68]If a person started a job in September 1, 2012, then tags are measured from August 1, 2011 to the day before August 1, 2012.

possible proxy for novel information is whether a new friend is formed through a pre-existing friend.[69] If this is true, friends-of-a-friend should have a lower correlation with job transmission than independently formed friendships. We include a dummy for friends-of-a-friend in our model and interact it with tags, posts and mutual friends. The results in Table 1.14 of the Appendix have insignificant coefficients, but they are of the signs we would expect. That is, a friend-of-a-friend is negatively correlated with job transmission. Also tags, posts and mutual friends from a non-friend-of-a-friend are generally more positively correlated with job transmission.[70] If more novel information flows through non-friends-of-a-friend, then this suggests that weak ties deliver more novel information.

1.5 Concluding Remarks

One of the most influential claims in the literature about social networks and labor markets has been Granovetter's "strength of weak ties" result which states that an individual is more likely to get a job with the help of a weak rather than strong tie.[71] We have decomposed that result into two different hypotheses based on the information that is available for analysis. We find, like the original paper, our data support the Descriptive Weak Ties Hypothesis, which concentrates on an individual and the specific friend(s) who helped her find a job. Over 90% of individuals were transmitted their most recent job from a weak tie. However, the distribution of tie strength in the population at large is also highly skewed toward weak ties, so this strength of weak ties is largely mechanical. We test our second hypothesis using the individual's full social

[69] Consider person A and her pre-existing friend B. Currently, A has access to information from B. If A makes a new friend, Z, who is also friends with B (B-Z already exists), then Z offers less new information than if Z was not already friends with B.

[70] We could only identify if a dyad was a friend-of-a-friend for 34,000 dyads which is only 2% of the alternative sub-sample. This is because Facebook only has the exact time of friendship formation for friendships formed after May 2006.

[71] For evidence of the influence of this paper refer to the citation count in footnote 3.

network and find increases in tie strength are associated with increases in the probability of job transmission from a friend. This relationship is not driven by individual-level unobservable variables or observable dyad-level variables, and remains after a number of robustness checks. In short, a person is most likely to be transmitted a job from a weak tie because weak ties are prolific in social networks. However, when a strong tie exists that strong tie has a higher probability of transmitting an individual a job.

Some theories about social networks and labor market outcomes have treated a tie as either existing or not existing, thus operationalizing weak or strong tie predictions through the structure of the network. Other theories have made assumptions that vary by the magnitude of tie strength between two individuals, which may be represented by contact or subjective self-reported feelings. Little work has used both these ideas of tie strength simultaneously. Network structure-based measures of tie strength (mutual friends and network overlap) are only weakly positively correlated with contact-based measures (tags, posts, and their scaled counterparts), and all these measures have power in predicting job transmission. This paper points out the importance of these different tie strength measures and suggests models incorporating these distinct tie strength measures will be an excellent area for future research.

In the future, we plan to use these data and methods to further investigate the relationship between job market outcomes and social ties. We have based our analysis on objective measures of tie strength, but a person's self-reported feelings about a friendship and how that friend helped in a job search are also important. We are in the process of running a survey with Facebook to collect data on these subjective feelings of tie strength and job search assistance, and we look forward to exploring the relationship between these measures and the results reported in this paper. The ultimate research goal suggested by our study is to find the causal effect of tie strength on job finding. We believe that both laboratory experiments and large-scale field experiments that exogenously affect the structure of the network or the level of contact between friends are a natural extension of our work.

Previous work has shown that a majority of jobs are found through social ties, and those who found a job via social ties have higher income, higher productivity, and longer tenure. This paper illustrates that whether strong or weak ties are more valuable in job search is a very nuanced question. The answer depends on the scope of the data used in the analysis. Contact and network structure-based measures need to be accounted for both empirically and theoretically. When looking at only data in which a person was helped by that specific friend, we find weak ties matter most. But when conditioning on information about all a person's social connections, strong ties are more influential than weak ties. In short, weak ties are important in aggregate because they are prolific, while strong ties are scarce but associated with a higher probability of job transmission.

1.6 Acknowledgement

Chapter 1, in part is currently being prepared for submission for publication of the material. Gee, Laura; Jones, Jason. The dissertation author was the primary investigator and author of this material.

1.7 Chapter 1 Appendix

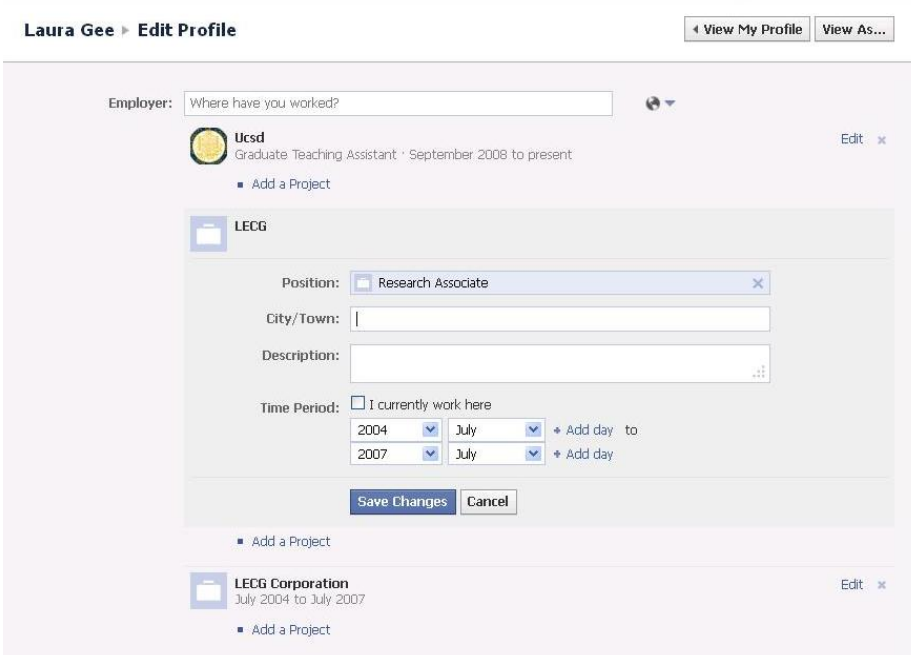


Figure 1.6: Employment Information Example

Note: An example of how a user can enter their employment information on Facebook.

Table 1.10: Primary Sub-sample Weighting Example

userid (i)	friendid (k)	J_{ik}	Orig. Weight	Over-Sample w. Prob.	New Weight
100	1	1	0.083	1	0.08
100	2	1	0.083	1	0.08
100	3	0	0.083	0.1	0.83
100	4	0	0.083	0.1	
100	5	0	0.083	0.1	
100	6	0	0.083	0.1	
100	7	0	0.083	0.1	
100	8	0	0.083	0.1	
100	9	0	0.083	0.1	
100	10	0	0.083	0.1	
100	11	0	0.083	0.1	
100	12	0	0.083	0.1	

1.7.1 Robustness Checks

Sensitivity Analysis

Table 1.11: Sensitivity to Controls in Linear Models Dependent Variable Job Transmission

	Controls				
	No Controls	Demographic	Col. 2 Plus Education	Col. 3 Plus Tenure	Col. 4 Plus Alter Friends
	1	**2**	**3**	**4**	**5**
			Panel A Tagging (10)		
Coefficient	0.009***	0.008***	0.008***	0.008***	0.008***
Standard Error	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
R2	0.073	0.075	0.078	0.121	0.121
			Panel B Posting (10)		
Coefficient	0.020***	0.019***	0.018***	0.019***	0.019***
Standard Error	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
R2	0.073	0.076	0.078	0.121	0.121
			Panel C Mutual Friends (100)		
Coefficient	-0.001***	-0.002***	0.004***	0.007***	0.007***
Standard Error	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
R2	0.073	0.075	0.077	0.121	0.121
N	1,017,089	1,017,089	1,017,089	1,017,089	1,017,089

All models include user-level fixed effect. Standard errors clustered at the user-level in parentheses. Each user's weights sum to 1. Legend: * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$
Controls in Col. 2: friend years older, both male/female, same relationship, same city/state
Controls added in Col. 3: friend more/less educated, same high school/college/grad. school
Controls added in Col. 4: friend's tenure at employer at user's start date
Controls added in Col. 5: alter friends (100), alter friends $(100)^2$, alter friends $(100)^3$, alter friends $(100)^4$

Employer Name and Industry

Industry classifications were done by matching the name of the employer listed in Facebook to names from the ReferenceUSA database available at www.referenceusa.com. We pulled the names and Primary Standard Industrial Classification (SIC) code and North American Industry Classification System (NAICS) codes for all US businesses with 500 or more employees, which was a total of 20,632 businesses with 16,788 unique company names.[72] For those businesses with multiple primary SIC or NAICS codes (this occurs because chain stores like Walmart are often listed multiple times with slight variations in the primary code), we kept the lowest SIC and lowest NAICS code per company name. This left us with a total of 16,788 businesses representing 2,212 SIC codes and 816 NAICS codes. We will primarily use the NAICS code because this is the current system used by Federal statistical agencies in classifying business establishments for the purpose of collecting, analyzing, and publishing statistical data related to the U.S. business economy. These were matched to the employer id numbers used by our Facebook dyads by the name listed on Facebook. To match these names to those in ReferenceUSA we first trimmed white space, converted all company names to all capital letters, removed punctuation (e.g. "." or "-"), articles (e.g. "THE" or "A"), common abbreviations (e.g. "U.S." converted to "UNITED STATES"), and removed common business titles (e.g. "CORP." or "INC."). We were able to match about 35% of the users to an industry, and only about 26% of their friends to an industry. Overall we matched both a user and friend's industry for only about 14% of the primary sub-sample, so we do not include this variable in the bulk of our analysis.

[72]See http://www.census.gov/epcd/www/sic.html and http://www.census.gov/eos/www/naics/

Table 1.12: Robustness Against Incidental Job Transmission

Variable	Bttm 75% Emplyrs	1 Mention Emplyrs	Same City Job	Distance	Coll. Users	Grad. Users	Coll. Dyads	Same Indust.	Indust. Dummies
	1	2	3	4	5	6	7	8	9
any dyad tag	0.016***	0.022***	0.014***	0.016***	0.014***	0.007***	0.015***	0.012***	0.015***
	(0.001)	(0.002)	(0.002)	(0.001)	(0.001)	(0.002)	(0.001)	(0.001)	(0.001)
tags (10)	0.003***	0.004**	0.005***	0.003***	0.002***	0.003*	0.002**	0.002***	0.002***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)	(0.001)	(0.001)	(0.001)
any dyad post	0.013***	0.014***	0.007***	0.012***	0.012***	0.010***	0.012***	0.011***	0.012***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
post (10)	0.007***	0.017***	0.005*	0.008***	0.008***	0.007**	0.008***	0.007***	0.008***
	(0.002)	(0.003)	(0.002)	(0.001)	(0.001)	(0.003)	(0.001)	(0.001)	(0.001)
any mutual friends	0.009***	0.009***	0.004	0.006***	0.005**	0.004	0.005*	0.004*	0.006***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.003)	(0.002)	(0.002)	(0.002)
mutual friends (100)	0.004***	0.006***	0.002***	0.005***	0.004***	0.002**	0.004***	0.004***	0.004***
	(0.000)	(0.001)	(0.000)	(0.000)	(0.000)	(0.001)	(0.000)	(0.000)	(0.000)
F years older (10)	-0.004***	-0.001	-0.004***	-0.009***	-0.008***	-0.006***	-0.009***	-0.008***	-0.008***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
both male	0.003***	0.004***	0.003***	0.008***	0.007***	0.006***	0.007***	0.005***	0.007***
	(0.000)	(0.001)	(0.001)	(0.000)	(0.000)	(0.001)	(0.000)	(0.000)	(0.000)
both female	0.005***	0.005***	0.002***	0.003***	0.003***	0.001	0.003***	0.003***	0.004***
	(0.000)	(0.001)	(0.001)	(0.000)	(0.000)	(0.001)	(0.000)	(0.000)	(0.000)
F more educated	0.002***	0.002	0.002*	0.001*	0.001		0.001	0.001	0.001
	(0.001)	(0.001)	(0.001)	(0.001)	(0.000)		(0.001)	(0.001)	(0.001)
F less educated	-0.008***	-0.006***	-0.008***	-0.008***	-0.008***	-0.012***	-0.012***	-0.006***	-0.007***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)

Includes user-level fixed effect. Standard errors clustered at the user-level. Weighted so that each user's weights sum to 1.

Legend: * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$ Continued on Next Page...

Table 1.12 – Continued

Variable	**Bttm 75% Emplyrs**	**1 Mention Emplyrs**	**Same City Job**	**Distance**	**Coll. Users**	**Grad. Users**	**Coll. Dyads**	**Same Indust.**	**Indust. Dummies**
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**
same relationship status	0.001*	0.002**	0.001**	0.001*	0.001	-0.000	0.000	0.001*	0.001*
	(0.000)	(0.001)	(0.000)	(0.000)	(0.000)	(0.001)	(0.000)	(0.000)	(0.000)
same state	0.007***	0.009***	0.003***		0.006***	0.009***	0.006***	0.005***	0.006***
	(0.000)	(0.001)	(0.000)		(0.000)	(0.001)	(0.000)	(0.000)	(0.000)
same city	0.011***	0.006***	0.077***		0.011***	0.013***	0.011***	0.007***	0.010***
	(0.001)	(0.001)	(0.001)		(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
F tenure	0.018***	0.016***	0.013***	0.022***	0.021***	0.019***	0.022***	0.020***	0.021***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
same high school	-0.017***	-0.014***	-0.008***	-0.021***	-0.020***	-0.022***	-0.021***	-0.016***	-0.020***
	(0.000)	(0.001)	(0.001)	(0.000)	(0.000)	(0.001)	(0.000)	(0.000)	(0.000)
same college	0.000	-0.002**	-0.005***	0.001*	0.000	-0.007***	0.000	-0.002***	-0.000
	(0.000)	(0.001)	(0.001)	(0.000)	(0.000)	(0.001)	(0.000)	(0.000)	(0.000)
same grad school	0.023***	0.014***	0.014***	0.024***	0.023***	0.019***	0.021***	0.014***	0.021***
	(0.002)	(0.003)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
alter friends (100)	-0.001**	-0.002***	-0.001**	-0.001***	-0.001*	0.000	-0.000	-0.001*	-0.001**
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.001)	(0.000)	(0.000)	(0.000)
alter friends $(100)^2$	0.000***	0.000***	0.000**	0.000***	0.000**	0.000	0.000	0.000**	0.000**
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
alter friends $(100)^3$	-0.000***	-0.000***	-0.000**	-0.000***	-0.000**	-0.000	-0.000*	-0.000**	-0.000***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
alter friends $(100)^4$	0.000***	0.000***	0.000*	0.000***	0.000**	0.000	0.000*	0.000**	0.000**
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
identified distance				0.010***					

Includes user-level fixed effect. Standard errors clustered at the user-level. Weighted so that each user's weights sum to 1.

Legend: * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$ Continued on Next Page…

Table 1.12 – Continued

Variable	Bttm 75% Emplyrs	1 Mention Emplyrs	Same City Job	Distance	Coll. Users	Grad. Users	Coll. Dyads	Same Indust.	Indust. Dummies
	1	2	3	4	5	6	7	8	9
				(0.000)					
distance (100km)				-0.001***					
				(0.000)					
identified ind.								-0.000	
								(0.000)	
same industry								0.271***	
								(0.002)	
F any industry									0.024***
									(0.002)
F Accommodation									0.000
									(0.003)
F Administrative									-0.000
									(0.005)
F Agriculture									0.027
									(0.022)
F Arts									-0.003
									(0.004)
F Construction									-0.012**
									(0.005)
F Educational									0.008***
									(0.002)
F Finance									-0.009***
									(0.002)

Includes user-level fixed effect. Standard errors clustered at the user-level. Weighted so that each user's weights sum to 1.

Legend: * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$ Continued on Next Page...

Table 1.12 – Continued

Variable	Bttm 75% Emplyrs	1 Mention Emplyrs	Same City Job	Distance	Coll. Users	Grad. Users	Coll. Dyads	Same Indust.	Indust. Dummies
	1	2	3	4	5	6	7	8	9
F Health Care									0.002
									(0.002)
F Information									0.010***
									(0.003)
F Management									-0.001
									(0.005)
F Manufacturing									-0.011***
									(0.003)
F Mining									-0.014*
									(0.005)
F Other Services									0.005
									(0.003)
F Professional									-0.003
									(0.002)
F Public Admin.									-0.003
									(0.003)
F Real Estate									-0.015*
									(0.006)
F Retail Trade									0.015***
									(0.002)
F Transportation									-0.012*
									(0.005)
F Wholesale									0.009**

Includes user-level fixed effect. Standard errors clustered at the user-level. Weighted so that each user's weights sum to 1.

Legend: * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$ Continued on Next Page...

Table 1.12 – Continued

Variable	Bttm 75% Emplyrs	1 Mention Emplyrs	Same City Job	Distance	Coll. Users	Grad. Users	Coll. Dyads	Same Indust.	Indust. Dummies
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**
									(0.003)
Constant	-0.014***	-0.014***	-0.022***	-0.006***	-0.006***	-0.002	-0.005**	-0.013***	-0.015***
	(0.002)	(0.003)	(0.002)	(0.002)	(0.002)	(0.004)	(0.002)	(0.002)	(0.002)
R2	0.095	0.085	0.115	0.122	0.118	0.120	0.200	0.200	0.128
N	760,156	277,865	369,802	1,017,089	953,428	213,724	865,089	1,017,089	1,017,089

Includes user-level fixed effect. Standard errors clustered at the user-level. Weighted so that each user's weights sum to 1.

Legend: * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$

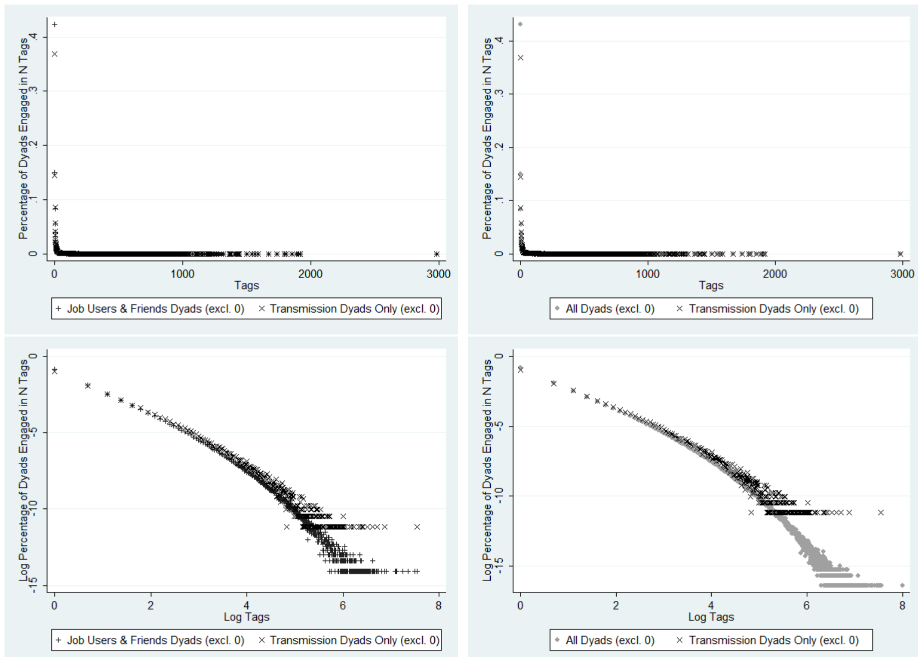


Figure 1.7: Distribution of Tags (excluding 0)

Note: The panels of this figure picture the distribution of weak to strong ties using number of tags (excluding dyads with 0 tags) from a user to a friend the year before the user began her most recent job as the measure of tie strength. The upper left panel shows the distribution for only those 63,000 dyads who had a job transmission (Transmission Dyads Only, black X) versus the 1.3 million dyads where the user of interest was transmitted her job from a friend (Job Users & Friends, dark grey +). The upper right panel shows the same distribution for only those 63,000 dyads who had a job transmission (black X) versus all the 13 million dyads with full information available in our data (All, light grey circle) with above zero tags. There are 63,000 dyads in the "Transmission Dyads Only" (black X) distribution. There are those same users connected to all their 1.3 million friends in the "Job Users & Friends" (dark grey +) distribution. There 13 million dyads in the "All" (light grey circle) distribution. The lower panels show the same information in log-log transformation.

Table 1.13: Robustness Against Different Tie and Tie Strength Definitions

Variable	**Contact Ties**	**User 2 Friend 2yr to 1yr**	**Friend 2 User**	**User 2 Friend**	**Friend 2 User 1yr 30day**	**Bi-Dir 1yr 30day**	**Bi-Dir 1yr 30day**
	1	**2**	**3**	**4**	**5**	**6**	**7**
any dyad tag	0.028***						
	(0.003)						
tagging (10)	0.006***						
	(0.002)						
any tag (2yr-1yr)		0.030***					
		(0.005)					
tags (2yr-1yr)		0.008***					
		(0.002)					
any dyad tag (F2U)			0.017***				
			(0.001)				
tags (10) (F2U)			0.003***				
			(0.001)				
any tag 1yr 30day (U2F)				0.015***			
				(0.001)			
tags 1yr 30day (10) (U2F)				0.002***			
				(0.001)			
any tag 1yr 30day (F2U)					0.016***		
					(0.001)		
tags 1yr 30day (10) (F2U)					0.004***		
					(0.001)		
any tag (bi-dir)						0.015***	
						(0.001)	

Includes user-level fixed effect. Standard errors clustered at the user-level. Weighted so that each user's weights sum to 1.

Legend: * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$ Continued on Next Page. . .

Table 1.13 – Continued

Variable	Contact Ties	User 2 Friend 2yr to 1yr	Friend 2 User	User 2 Friend	Friend 2 User 1yr 30day	Bi-Dir 1yr 30day	Bi-Dir 1yr 30day
	1	2	3	4	5	6	7
tags (10) (bi-dir)						0.003***	
						(0.001)	
any tag 1yr 30day (bi-dir)							0.015***
							(0.001)
tags 1yr 30day (10) (bi-dir)							0.000***
							(0.000)
any dyad post	0.002		0.012***	0.012***	0.012***	0.011***	0.011***
	(0.004)		(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
posting (10)	0.014***		0.007***	0.008***	0.007***	0.006***	0.006***
	(0.003)		(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
any mutual friends	-0.005		0.006***	0.006***	0.006***	0.006***	0.006***
	(0.009)		(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
mutual friends (100)	0.005*		0.004***	0.005***	0.005***	0.004***	0.004***
	(0.002)		(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
F years older (10)	-0.021***	-0.018***	-0.009***	-0.009***	-0.009***	-0.009***	-0.009***
	(0.002)	(0.005)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
both male	0.026***	0.032***	0.008***	0.008***	0.008***	0.008***	0.008***
	(0.003)	(0.005)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
both female	0.011***	0.012**	0.003***	0.004***	0.003***	0.003***	0.003***
	(0.003)	(0.004)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
F more educated	0.000	-0.001	0.001**	0.001**	0.001**	0.001**	0.001**
	(0.003)	(0.005)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
F less educated	-0.025***	-0.024***	-0.008***	-0.008***	-0.008***	-0.008***	-0.008***

Includes user-level fixed effect. Standard errors clustered at the user-level. Weighted so that each user's weights sum to 1.

Legend: * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$ Continued on Next Page. . .

Table 1.13 – Continued

Variable	**Contact Ties**	**User 2 Friend 2yr to 1yr**	**Friend 2 User**	**User 2 Friend**	**Friend 2 User 1yr 30day**	**Bi-Dir 1yr 30day**	**Bi-Dir 1yr 30day**
	1	**2**	**3**	**4**	**5**	**6**	**7**
	(0.003)	(0.006)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
same relationship status	0.004	0.008*	0.001*	0.001*	0.001*	0.001*	0.001*
	(0.002)	(0.004)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
same state at start date	0.015***	0.006	0.006***	0.006***	0.006***	0.006***	0.006***
	(0.003)	(0.004)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
same city at start date	0.025***	0.038***	0.010***	0.010***	0.010***	0.010***	0.010***
	(0.003)	(0.005)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
F tenure at employer	0.056***	0.063***	0.022***	0.022***	0.022***	0.022***	0.022***
	(0.001)	(0.001)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
same high school	-0.047***	-0.036***	-0.020***	-0.020***	-0.020***	-0.020***	-0.020***
	(0.003)	(0.004)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
same college	0.006*	0.018***	-0.000	0.000	0.000	-0.000	-0.000
	(0.003)	(0.004)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
same grad school	0.071***	0.096***	0.023***	0.023***	0.023***	0.023***	0.023***
	(0.009)	(0.014)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
alter friends (100)	0.002	0.003	-0.001***	-0.001***	-0.001***	-0.001**	-0.001**
	(0.002)	(0.003)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
alter friends $(100)^2$	-0.000	-0.000	0.000***	0.000***	0.000***	0.000***	0.000***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
alter friends $(100)^3$	0.000	0.000	-0.000***	-0.000***	-0.000***	-0.000***	-0.000***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
alter friends $(100)^4$	-0.000	-0.000	0.000***	0.000***	0.000***	0.000***	0.000***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)

Includes user-level fixed effect. Standard errors clustered at the user-level. Weighted so that each user's weights sum to 1.

Legend: * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$ Continued on Next Page. . .

Table 1.13 – Continued

Variable	**Contact Ties**	**User 2 Friend 2yr to 1yr**	**Friend 2 User**	**User 2 Friend**	**Friend 2 User 1yr 30day**	**Bi-Dir 1yr 30day**	**Bi-Dir 1yr 30day**
	1	**2**	**3**	**4**	**5**	**6**	**7**
Constant	0.114***	0.106***	-0.008***	-0.008***	-0.008***	-0.008***	-0.008***
	(0.010)	(0.008)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
R2	0.762	0.799	0.123	0.122	0.122	0.123	0.123
N	131,183	61,630	1,017,109	1,017,109	1,017,109	1,017,109	1,017,109

Includes user-level fixed effect. Standard errors clustered at the user-level. Weighted so that each user's weights sum to 1.

Legend: * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$

Novel Information: Friends of Friends

Facebook has the exact time a friendship was formed for friendships formed after May 2006. We can identify friendships through a mutual friend, but it is very noisy because real-life friends are not always Facebook friends and often they become Facebook friends well after the real life friendship was formed. We identify if a friendship was formed through a mutual friend by comparing the times that the friendships began on Facebook.[73]

Table 1.14: Friends of Friends

Variable	FOF	Tags	Posts	Mfriends	All
Friend-of-Friend (FOF)	-0.004	-0.003	-0.003	-0.007	-0.007
	(0.005)	(0.005)	(0.005)	(0.007)	(0.007)
tagging (10)		0.009			0.008
		(0.008)			(0.008)
tagging (10)*FOF		-0.000			-0.003
		(0.006)			(0.006)
posting (10)			0.016		0.012
			(0.014)		(0.013)
dyad posting (10)*FOF			0.008		0.008
			(0.016)		(0.015)
mutual friends (100)				0.011	0.006
				(0.008)	(0.008)
mutual friends (100)*FOF				0.004	0.008
				(0.007)	(0.008)
F years older (10)	-0.016**	-0.016**	-0.016**	-0.015**	-0.015**
	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)
both male	0.004	0.004	0.004	0.003	0.003
	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)
both female	0.005	0.005	0.005	0.005	0.004
	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)
F more educated	0.004	0.004	0.004	0.005	0.005

Includes user-level fixed effect. Standard errors clustered at the user-level. Weighted so that each user's weights sum to 1

Legend: * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$ Continued on Next Page...

[73]Person i can become friends with person k through person j in two ways. The first way is if person i and j become friends (t_{ij}), then person j and k become friends (t_{jk}), and last person i and k become friends (t_{ik}). The second way is if person j and k become friends (t_{jk}), then person i and j become friends (t_{ij}), and last person i and k become friends (t_{jk}). If we find either $t_{ik} > t_{jk} > t_{ij}$ or $t_{ik} > t_{ij} > t_{jk}$ then the dyad ik is a friend-of-a-friend ($FOF_{ij} = 1$ if friends of friends).

Table 1.14 – Continued

Variable	FOF	Tags	Posts	Mfriends	All
	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)
F less educated	0.006	0.006	0.007	0.006	0.006
	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)
same relationship status	0.005	0.005	0.005	0.005	0.004
	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)
same state at start date	0.006	0.006	0.006	0.005	0.005
	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)
same city at start date	0.011	0.011	0.011	0.010	0.010
	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)
F tenure at employer	0.035***	0.035***	0.035***	0.036***	0.035***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
same high school	-0.012*	-0.012*	-0.012*	-0.017**	-0.016**
	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)
same college	0.026***	0.025***	0.026***	0.026***	0.025***
	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)
same grad school	-0.081	-0.082	-0.082	-0.084	-0.084
	(0.089)	(0.089)	(0.089)	(0.089)	(0.089)
alter friends (100)	-0.005	-0.005	-0.005	-0.007*	-0.007*
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
alter friends $(100)^2$	0.001	0.001	0.001	0.001	0.001
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
alter friends $(100)^3$	-0.000	-0.000	-0.000	-0.000	-0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
alter friends $(100)^4$	0.000	0.000	0.000	0.000	0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
constant	0.019*	0.019*	0.018*	0.021*	0.021*
	(0.008)	(0.008)	(0.009)	(0.009)	(0.009)
r2	0.160	0.161	0.161	0.161	0.162
N	34,783	34,783	34,783	34,783	34,783

Includes user-level fixed effect. Standard errors clustered at the user-level. Weighted so that each user's weights sum to 1. Legend: * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$

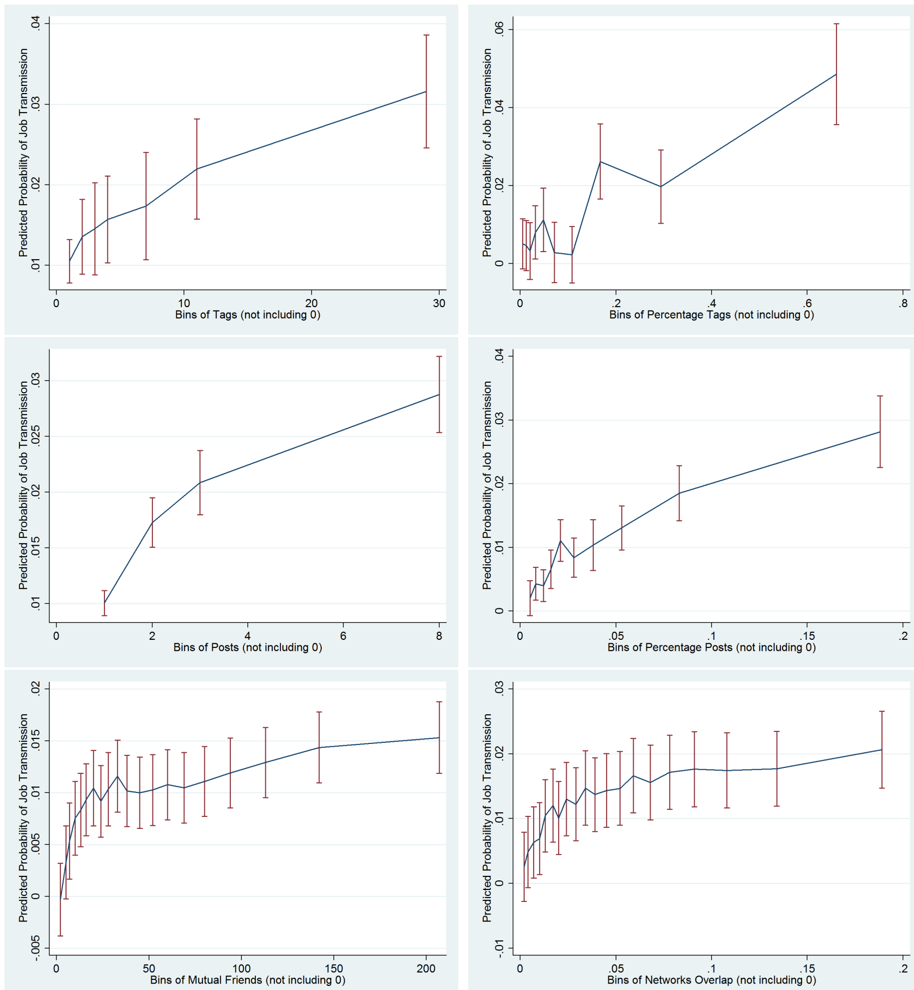


Figure 1.8: Non-parametric Plots

Note: The panels of this figure plot the predicted probability of job transmission by quantile for each of our tie strength measures. Each coefficient is plotted at the median level of the tie strength variable in that quantile and the 95% confidence interval is shown around each coefficient. On the left hand side of this Figure are the raw tie strength measures from the primary sub-sample. On the right hand side of this figure are the scaled tie strength measures from the alternative sub-sample. The results for the underlying regressions are available from the authors by request.

Table 1.15: Linear Models Dependent Variable Job Transmission All Tie Strength Measures (Replication of Results in Table 1.9 in Text Using Other Sub-samples for Each Model)

Variable	**Exten.**	**Inten.**	**Ext./Int.**	**Inten.**	**Ext./Int.**
		Absolute	**Absolute**	**Pct.**	**Pct.**
	Alt. Sample	**Alt. Sample**	**Alt. Sample**	**Primary Sample**	**Primary Sample**
	1	**2**	**3**	**4**	**5**
any dyad tag	0.019***		0.015***		-0.059***
	(0.002)		(0.002)		(0.001)
any dyad post	0.014***		0.012***		-0.032***
	(0.001)		(0.001)		(0.000)
any mutual friends	0.011***		0.010***		0.004**
	(0.003)		(0.003)		(0.002)
tags (10)		0.005***	0.003**		
		(0.001)	(0.001)		
posts (10)		0.016***	0.008***		
		(0.002)	(0.002)		
mutual friends (100)		0.007***	0.005***		
		(0.001)	(0.001)		
% tags				0.600***	0.759***
				(0.007)	(0.008)
% posts				0.705***	0.826***
				(0.006)	(0.006)
network overlap				0.003	0.044***
				(0.004)	(0.004)
F years older (10)	-0.009***	-0.009***	-0.009***	-0.009***	-0.009***

Includes user-level fixed effect. Standard errors clustered at the user-level. Weighted so that each user's weights sum to 1.

Legend: * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$ Continued on Next Page...

Table 1.15 – Continued

Variable	Exten.	Inten. Absolute	Ext./Int. Absolute	Inten. Pct.	Ext./Int. Pct.
	Alt. Sample	Alt. Sample	Alt. Sample	Primary Sample	Primary Sample
	1	**2**	**3**	**4**	**5**
	(0.001)	(0.001)	(0.001)	(0.001)	(0.000)
both male	0.009***	0.009***	0.009***	0.007***	0.007***
	(0.001)	(0.001)	(0.001)	(0.000)	(0.000)
both female	0.004***	0.004***	0.004***	0.002***	0.003***
	(0.001)	(0.001)	(0.001)	(0.000)	(0.000)
F more educated	0.001	0.002*	0.002*	0.002**	0.002**
	(0.001)	(0.001)	(0.001)	(0.000)	(0.000)
F less educated	-0.007***	-0.007***	-0.007***	-0.007***	-0.008***
	(0.001)	(0.001)	(0.001)	(0.000)	(0.000)
same relationship status	0.001**	0.001**	0.001*	-0.000	0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
same state at start date	0.006***	0.006***	0.006***	0.005***	0.005***
	(0.001)	(0.001)	(0.001)	(0.000)	(0.000)
same city at start date	0.009***	0.010***	0.009***	0.008***	0.010***
	(0.001)	(0.001)	(0.001)	(0.000)	(0.000)
F tenure at employer	0.022***	0.022***	0.022***	0.021***	0.020***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
same high school	-0.020***	-0.023***	-0.022***	-0.017***	-0.020***
	(0.001)	(0.001)	(0.001)	(0.000)	(0.000)
same college	0.001	0.001	0.001	-0.001	0.000
	(0.001)	(0.001)	(0.001)	(0.000)	(0.000)
same grad school	0.028***	0.028***	0.028***	0.020***	0.021***

Includes user-level fixed effect. Standard errors clustered at the user-level. Weighted so that each user's weights sum to 1.

Legend: * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$ Continued on Next Page...

Table 1.15 – Continued

Variable	**Exten.**	**Inten.**	**Ext./Int.**	**Inten.**	**Ext./Int.**
		Absolute	**Absolute**	**Pct.**	**Pct.**
	Alt. Sample	**Alt. Sample**	**Alt. Sample**	**Primary Sample**	**Primary Sample**
	1	**2**	**3**	**4**	**5**
	(0.003)	(0.003)	(0.003)	(0.002)	(0.002)
alter friends (100)	-0.001	-0.001***	-0.001***	0.001*	-0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
alter friends $(100)^2$	0.000**	0.000***	0.000***	-0.000	0.000*
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
alter friends $(100)^3$	-0.000***	-0.000***	-0.000***	-0.000	-0.000**
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
alter friends $(100)^4$	0.000***	0.000***	0.000***	0.000	0.000**
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Constant	-0.010***	0.004**	-0.009**	-0.007***	-0.004*
	(0.003)	(0.001)	(0.003)	(0.001)	(0.002)
R2	0.132	0.132	0.132	0.181	0.189
N	1,438,699	1,438,699	1,438,699	1,017,089	1,017,089

Includes user-level fixed effect. Standard errors clustered at the user-level. Weighted so that each user's weights sum to 1.

Legend: * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$

Table 1.16: Conditional Logit Models Dependent Variable Job Transmission (Alternative Sub-sample)

Variable	**Alternative Abs.**	**Alternative Pct.**
	1	**2**
any dyad tag	0.422***	0.287***
	(0.039)	(0.044)
tagging (10)	0.042***	
	(0.011)	
any dyad post	0.461***	0.417***
	(0.026)	(0.027)
posting (10)	0.127***	
	(0.034)	
any mutual friends	0.326***	
	(0.090)	
mutual friends (100)	0.198***	
	(0.024)	
% tags		0.601***
		(0.129)
% posts		0.506**
		(0.162)
any network overlap		0.277**
		(0.090)
network overlap		2.245***
		(0.240)
F years older (10)	-0.323***	-0.318***
	(0.025)	(0.025)
both male	0.326***	0.320***
	(0.028)	(0.028)
both female	0.167***	0.170***
	(0.028)	(0.027)
F more educated	0.039	0.047
	(0.029)	(0.029)
F less educated	-0.232***	-0.231***
	(0.034)	(0.034)
same relationship status	0.051*	0.047*
	(0.020)	(0.020)
same state at start date	0.259***	0.255***

Includes user-level fixed effect.

Standard errors clustered at the user-level.

Weighted so that each user's weights sum to 1.

Legend: * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$

Continued on Next Page…

Table 1.16 – Continued

Variable	**Alternative Abs. 1**	**Alternative Pct. 2**
	(0.029)	(0.029)
same city at start date	0.329***	0.325***
	(0.030)	(0.030)
F tenure at employer	0.602***	0.602***
	(0.005)	(0.005)
same high school	-0.943***	-0.984***
	(0.031)	(0.032)
same college	0.144***	0.142***
	(0.029)	(0.029)
same grad school	0.888***	0.882***
	(0.083)	(0.084)
alter friends (100)	-0.051***	-0.041**
	(0.014)	(0.014)
alter friends $(100)^2$	0.006***	0.006***
	(0.002)	(0.002)
alter friends $(100)^3$	-0.000***	-0.000***
	(0.000)	(0.000)
alter friends $(100)^4$	0.000***	0.000***
	(0.000)	(0.000)
Psuedo R2	0.200	0.201
N	1,438,699	1,438,699

Includes user-level fixed effect.

Standard errors clustered at the user-level.

Weighted so that each user's weights sum to 1.

Legend: * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$

Alternative Sub-sample Summary Statistics

In the following we present data for a random sample of 12,000 users and all their friends which we call the Alternative Sub-sample. Dyads with $J_{ik} = 1$ and $J_{ik} = 0$ are all sampled with 100% probability conditional on a user i being in our randomly selected sub-sample. Recall that in the primary sub-sample the $J_{ik} = 1$ dyads are over-sampled as compared to the $J_{ik} = 0$ dyads.

Table 1.17: User statistics (Alternative Sub-sample)

Variable	Mean	Std. Dev.	Min.	Max.
age	24.58	5.418	16	65
male	0.475	0.499	0	1
some high school	0.091	0.287	0	1
some college	0.719	0.45	0	1
some post BA	0.191	0.393	0	1
married	0.179	0.384	0	1
single	0.293	0.455	0	1
in relationship	0.259	0.438	0	1
engaged	0.058	0.233	0	1
friend count	521.039	345.134	0	4653

N=1,438,699 Dyads
12,263 Users
Dyad-level observation weighted
by 1/(number of times user is in data)

1.7.2 In Text Tables Details & Complimentary Tables

Table 1.18: Friend statistics (Alternative Sub-sample)

Variable	Mean	Std. Dev.	Min.	Max.
F age	25.029	5.516	16	65
F male	0.472	0.499	0	1
F married	0.202	0.402	0	1
F single	0.277	0.448	0	1
F in relationship	0.254	0.435	0	1
F engaged	0.06	0.237	0	1
F some high school	0.089	0.285	0	1
F some college	0.727	0.446	0	1
F some post BA	0.184	0.388	0	1
F friend count	499.385	380.165	0	6329
N=1,438,699 Dyads				
1,149,562 Friends				
Dyad-level observation weighted				
by 1/(number of times friend is in data)				

Table 1.19: Dyad-level Tie Strength Summary Statistics (Alternative Sub-sample)

Variable	Mean	Std. Dev.	Min.	Max.	N
job transmitted	0.02	0.141	0	1	1438699
any dyad tag	0.042	0.201	0	1	1438699
tags	0.28	3.744	0	974	1438699
tags (1+)	6.621	17.002	1	974	60919
% of tags	0.006	0.053	0	1	1438699
% of tags (1+)	0.147	0.217	0.001	1	60919
any dyad post	0.161	0.368	0	1	1438699
posts	0.366	2.19	0	627	1438699
posts (1+)	2.268	5.039	1	627	60919
% of posts	0.008	0.039	0	1	1438699
% of posts (1+)	0.051	0.086	0.001	1	60919
mutual friends	54.957	58.792	0	1303	1438699
any network overlap	0.986	0.117	0	1	1438699
network overlap	0.051	0.051	0	1	1438699
1,438,699 Dyads					
No Weighting					

Table 1.20: Correlation Absolute Tie Strength (Alternative Sub-sample)

Variables	mutual friends	total dyad tags	total dyad posts
mutual friends	1.000		
total dyad tags	0.036	1.000	
total dyad posts	0.039	0.333	1.000

Table 1.21: Dyad-level Demographic Summary Statistics

Variable	**Mean**	**Std. Dev.**	**Min.**	**Max.**	**N**
F years older (10)	0.044	0.453	-4.4	4.8	1438699
both male	0.251	0.434	0	1	1438699
both female	0.289	0.453	0	1	1438699
F more educated	0.153	0.36	0	1	1438699
F less educated	0.192	0.394	0	1	1438699
F more friends (100)	-1.479	5.464	-46.53	55.92	1438699
same relationship status	0.283	0.45	0	1	1438699
same state at start date	0.477	0.499	0	1	1438699
same city at start date	0.177	0.382	0	1	1438699
F tenure at employer (years)	1.121	1.773	-4.088	42.027	1438699
same high school	0.293	0.455	0	1	1438699
same college	0.308	0.462	0	1	1438699
same grad school	0.016	0.126	0	1	1438699
identified distance	0.322	0.467	0	1	1438699
min. distance	244.469	580.008	0	4649.6	463446
identified industries	0.013	0.114	0	1	1438699
same industry (NAICS)	0.999	0.034	0	1	18993
1,017,089 Dyads					
Weighted by inverse of sampling probability					

Table 1.22: Absolute Tie Strength Measures One-by-One, Full Results of Table 1.7 in Text

Variable	**Tag**	**Post**	**M Frnd**	**Tag**	**Post**	**M Frnd**	**Tag**	**Post**	**M Frnd**
	Exten.	**Exten.**	**Exten.**	**Inten.**	**Inten.**	**Inten.**	**Ext/Int**	**Ext/Int**	**Ext/Int**
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**
any dyad tag	0.025***						0.022***		
	(0.001)						(0.001)		
any dyad post		0.017***						0.014***	
		(0.000)						(0.001)	
any mutual friends			0.008***						0.007***
			(0.002)						(0.002)
tags (10)				0.008***			0.004***		
				(0.001)			(0.001)		
posts (10)					0.019***			0.012***	
					(0.002)			(0.002)	
mutual friends (100)						0.007***			0.007***
						(0.000)			(0.000)
F years older (10)	-0.009***	-0.010***	-0.010***	-0.009***	-0.009***	-0.009***	-0.009***	-0.010***	-0.009***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
both male	0.008***	0.008***	0.008***	0.008***	0.008***	0.008***	0.008***	0.008***	0.008***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
both female	0.004***	0.004***	0.004***	0.004***	0.004***	0.004***	0.004***	0.004***	0.004***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
F more educated	0.001*	0.001*	0.001*	0.001*	0.001*	0.001*	0.001*	0.001*	0.001*
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
F less educated	-0.008***	-0.008***	-0.008***	-0.008***	-0.008***	-0.008***	-0.008***	-0.008***	-0.008***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)

Includes user-level fixed effect. Standard errors clustered at the user-level. Weighted so that each user's weights sum to 1.

Legend: * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$ Continued on Next Page…

Table 1.22 – Continued

Variable	Tag	Post	M Frnd	Tag	Post	M Frnd	Tag	Post	M Frnd
	Exten.	**Exten.**	**Exten.**	**Inten.**	**Inten.**	**Inten.**	**Ext/Int**	**Ext/Int**	**Ext/Int**
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**
same relationship status	0.001**	0.001**	0.001**	0.001**	0.001**	0.001**	0.001**	0.001**	0.001**
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
same state at start date	0.006***	0.006***	0.006***	0.006***	0.006***	0.006***	0.006***	0.006***	0.006***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
same city at start date	0.011***	0.011***	0.011***	0.011***	0.011***	0.011***	0.011***	0.011***	0.011***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
F tenure at employer	0.022***	0.022***	0.022***	0.022***	0.022***	0.022***	0.022***	0.022***	0.022***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
same high school	-0.019***	-0.018***	-0.019***	-0.019***	-0.019***	-0.022***	-0.019***	-0.018***	-0.022***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
same college	0.001	0.001	0.001**	0.001*	0.001*	0.001	0.001	0.001	0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
same grad school	0.024***	0.024***	0.025***	0.025***	0.025***	0.024***	0.024***	0.024***	0.024***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
alter friends (100)	-0.000	-0.000	-0.001**	-0.001*	-0.001*	-0.001***	-0.000	-0.000	-0.001***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
alter friends $(100)^2$	0.000**	0.000*	0.000***	0.000**	0.000**	0.000***	0.000**	0.000*	0.000***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
alter friends $(100)^3$	-0.000***	-0.000**	-0.000***	-0.000***	-0.000***	-0.000***	-0.000***	-0.000**	-0.000***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
alter friends $(100)^4$	0.000**	0.000**	0.000***	0.000***	0.000**	0.000***	0.000**	0.000**	0.000***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
constant	0.000	-0.002**	-0.006***	0.001	0.001	0.002*	0.000	-0.002**	-0.004**

Includes user-level fixed effect. Standard errors clustered at the user-level. Weighted so that each user's weights sum to 1.

Legend: * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$ Continued on Next Page. . .

Table 1.22 – Continued

Variable	**Tag**	**Post**	**M Frnd**	**Tag**	**Post**	**M Frnd**	**Tag**	**Post**	**M Frnd**
	Exten.	**Exten.**	**Exten.**	**Inten.**	**Inten.**	**Inten.**	**Ext/Int**	**Ext/Int**	**Ext/Int**
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**
	(0.001)	(0.001)	(0.002)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)
R2	0.121	0.122	0.120	0.121	0.121	0.121	0.121	0.122	0.121
N	1,017,089	1,017,089	1,017,089	1,017,089	1,017,089	1,017,089	1,017,089	1,017,089	1,017,089

Includes user-level fixed effect. Standard errors clustered at the user-level. Weighted so that each user's weights sum to 1.

Legend: * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$

Table 1.23: Linear Models Dependent Variable Job Transmission Scaled Tie Strength Measures One-by-One (Alternative Sample), Full Results of Table 1.8 in Text

Variable	Inten. Pct. Tag 1	Inten. Pct. Post 2	Inten. Overlap 3	Ext/Int Pct. Tag 4	Ext/Int Pct. Post 5	Ext/Int Overlap 6
any dyad tag				0.011***		
				(0.002)		
any dyad post					0.009***	
					(0.001)	
any mutual friends						0.009***
						(0.003)
% of tags	0.094***			0.074***		
	(0.008)			(0.010)		
% of posts		0.124***			0.093***	
		(0.011)			(0.013)	
network overlap			0.095***			0.093***
			(0.007)			(0.007)
F years older (10)	-0.009***	-0.009***	-0.009***	-0.009***	-0.009***	-0.009***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
both male	0.009***	0.009***	0.009***	0.009***	0.009***	0.009***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
both female	0.005***	0.004***	0.005***	0.005***	0.004***	0.005***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
F more educated	0.001	0.001	0.002*	0.001	0.001	0.002*
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)

Includes user-level fixed effect. Standard errors clustered at the user-level. Weighted so that each user's weights sum to 1.

Legend: * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$ Continued on Next Page...

Table 1.23 – Continued

Variable	Inten. Pct. Tag 1	Inten. Pct. Post 2	Inten. Overlap 3	Ext/Int Pct. Tag 4	Ext/Int Pct. Post 5	Ext/Int Overlap 6
F less educated	-0.007***	-0.007***	-0.007***	-0.007***	-0.007***	-0.007***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
same relationship status	0.001**	0.001**	0.001**	0.001**	0.001**	0.001**
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
same state at start date	0.006***	0.007***	0.006***	0.006***	0.006***	0.006***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
same city at start date	0.009***	0.010***	0.010***	0.009***	0.010***	0.010***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
F tenure at employer	0.022***	0.022***	0.022***	0.022***	0.022***	0.022***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
same high school	-0.020***	-0.020***	-0.025***	-0.020***	-0.020***	-0.025***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
same college	0.002*	0.002**	0.001	0.002*	0.002*	0.001
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
same grad school	0.029***	0.029***	0.028***	0.029***	0.029***	0.028***
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
alter friends (100)	-0.001*	-0.001	-0.002***	-0.001	-0.001	-0.002***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
alter friends $(100)^2$	0.000**	0.000**	0.000***	0.000**	0.000**	0.000***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
alter friends $(100)^3$	-0.000***	-0.000***	-0.000***	-0.000***	-0.000***	-0.000***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
alter friends $(100)^4$	0.000***	0.000***	0.000***	0.000***	0.000***	0.000***

Includes user-level fixed effect. Standard errors clustered at the user-level. Weighted so that each user's weights sum to 1.

Legend: * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$ Continued on Next Page...

Table 1.23 – Continued

Variable	**Inten. Pct. Tag 1**	**Inten. Pct. Post 2**	**Inten. Overlap 3**	**Ext/Int Pct. Tag 4**	**Ext/Int Pct. Post 5**	**Ext/Int Overlap 6**
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
constant	0.003**	0.002	0.002	0.003*	0.001	-0.006*
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.003)
R2	0.132	0.132	0.131	0.132	0.132	0.131
N	1438699.000	1,438,699	1,438,699	1,438,699	1,438,699	1,438,699

Includes user-level fixed effect. Standard errors clustered at the user-level. Weighted so that each user's weights sum to 1.

Legend: * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$

Table 1.24: Scaled Tie Strength Measures One-by-One (Primary Sub-sample replication of Table 1.8 in Text)

Variable	**Pct Tag**	**Pct Post**	**Overlap**	**Pct Tag**	**Pct Post**	**Overlap**
	Inten.	**Inten.**	**Inten.**	**Ext/Int**	**Ext/Int**	**Ext/Int**
	1	**2**	**3**	**4**	**5**	**6**
any dyad tag				-0.068***		
				(0.001)		
any dyad post					-0.038***	
					(0.000)	
any mutual friends						0.005**
						(0.002)
% of tags	0.824***			1.009***		
	(0.004)			(0.003)		
% of posts		0.848***			0.990***	
		(0.004)			(0.004)	
network overlap			0.085***			0.083***
			(0.004)			(0.004)
F years older (10)	-0.009***	-0.009***	-0.009***	-0.009***	-0.009***	-0.009***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
both male	0.007***	0.008***	0.008***	0.008***	0.008***	0.008***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
both female	0.003***	0.002***	0.004***	0.004***	0.003***	0.004***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
F more educated	0.001**	0.001**	0.001**	0.001*	0.001**	0.001**
	(0.001)	(0.000)	(0.001)	(0.001)	(0.000)	(0.001)
F less educated	-0.008***	-0.007***	-0.008***	-0.008***	-0.008***	-0.008***
	(0.001)	(0.000)	(0.001)	(0.001)	(0.000)	(0.001)

Includes user-level fixed effect. Standard errors clustered at the user-level. Weighted so that each user's weights sum to 1.

Legend: * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$ Continued on Next Page...

Table 1.24 – Continued

Variable	Pct Tag Inten. 1	Pct Post Inten. 2	Overlap Inten. 3	Pct Tag Ext/Int 4	Pct Post Ext/Int 5	Overlap Ext/Int 6
same relationship status	0.000	0.000	0.001**	0.001	0.001	0.001**
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
same state at start date	0.005***	0.005***	0.006***	0.006***	0.006***	0.006***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
same city at start date	0.009***	0.009***	0.011***	0.010***	0.010***	0.011***
	(0.001)	(0.000)	(0.001)	(0.001)	(0.000)	(0.001)
F tenure at employer	0.021***	0.021***	0.022***	0.021***	0.021***	0.022***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
same high school	-0.018***	-0.017***	-0.023***	-0.019***	-0.018***	-0.023***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
same college	-0.000	0.000	0.000	0.001	0.001*	0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
same grad school	0.022***	0.022***	0.023***	0.023***	0.023***	0.023***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
alter friends (100)	-0.000	0.000	-0.001***	-0.000	0.000	-0.001***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
alter friends $(100)^2$	0.000	-0.000	0.000***	0.000*	0.000	0.000***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
alter friends $(100)^3$	-0.000*	-0.000	-0.000***	-0.000**	-0.000	-0.000***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
alter friends $(100)^4$	0.000*	0.000	0.000***	0.000**	0.000	0.000***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
constant	-0.001	-0.007***	-0.000	0.001*	-0.000	-0.005**

Includes user-level fixed effect. Standard errors clustered at the user-level. Weighted so that each user's weights sum to 1.

Legend: * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$ Continued on Next Page...

Table 1.24 – Continued

Variable	**Pct Tag Inten. 1**	**Pct Post Inten. 2**	**Overlap Inten. 3**	**Pct Tag Ext/Int 4**	**Pct Post Ext/Int 5**	**Overlap Ext/Int 6**
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)
R2	0.152	0.165	0.121	0.157	0.170	0.121
N	1,017,089	1,017,089	1,017,089	1,017,089	1,017,089	1,017,089

Includes user-level fixed effect. Standard errors clustered at the user-level. Weighted so that each user's weights sum to 1.

Legend: * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$

Table 1.25: Correlation Percentage Tie Strength (Primary Sub-sample replication of Table 1.5 in Text)

Variables	network overlap	percentage tags	percentage posts
network overlap	1.000		
percentage tags	0.057	1.000	
percentage posts	0.046	0.336	1.000

Table 1.26: Linear Models Dependent Variable Job Transmission All Tie Strength Measures, Full Results of Table 1.9 in Text

Variable	**Exten.**	**Inten.**	**Ext./Int.**	**Inten.**	**Ext./Int.**
		Absolute	**Absolute**	**Pct.**	**Pct.**
	Primary Sample	**Primary Sample**	**Primary Sample**	**Alt. Sample**	**Alt. Sample**
	1	**2**	**3**	**4**	**5**
any dyad tag	0.019***		0.015***		0.005**
	(0.001)		(0.001)		(0.002)
any dyad post	0.014***		0.012***		0.007***
	(0.000)		(0.001)		(0.001)
any mutual friends	0.007***		0.006***		0.009**
	(0.002)		(0.002)		(0.003)
tagging (10)		0.005***	0.002***		
		(0.001)	(0.001)		
posts (10)		0.016***	0.008***		
		(0.002)	(0.001)		
mutual friends (100)		0.006***	0.005***		
		(0.000)	(0.000)		
% of tags				0.069***	0.058***
				(0.008)	(0.010)
% of posts				0.098***	0.072***
				(0.011)	(0.013)
network overlap				0.075***	0.068***
				(0.007)	(0.007)
F years older (10)	-0.010***	-0.009***	-0.009***	-0.009***	-0.009***

Includes user-level fixed effect. Standard errors clustered at the user-level. Weighted so that each user's weights sum to 1.

Legend: * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$ Continued on Next Page...

Table 1.26 – Continued

Variable	**Exten.**	**Inten.**	**Ext./Int.**	**Inten.**	**Ext./Int.**
		Absolute	**Absolute**	**Pct.**	**Pct.**
	Primary Sample	**Primary Sample**	**Primary Sample**	**Alt. Sample**	**Alt. Sample**
	1	**2**	**3**	**4**	**5**
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
both male	0.008***	0.008***	0.008***	0.009***	0.009***
	(0.000)	(0.000)	(0.000)	(0.001)	(0.001)
both female	0.004***	0.004***	0.003***	0.004***	0.004***
	(0.000)	(0.000)	(0.000)	(0.001)	(0.001)
F more educated	0.001*	0.001*	0.001**	0.002*	0.002*
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
F less educated	-0.008***	-0.008***	-0.008***	-0.007***	-0.007***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
same relationship status	0.001*	0.001*	0.001*	0.001*	0.001*
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
same state at start date	0.006***	0.006***	0.006***	0.006***	0.006***
	(0.000)	(0.000)	(0.000)	(0.001)	(0.001)
same city at start date	0.010***	0.011***	0.010***	0.009***	0.009***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
F tenure at employer	0.022***	0.022***	0.022***	0.022***	0.022***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
same high school	-0.019***	-0.021***	-0.020***	-0.024***	-0.023***
	(0.000)	(0.000)	(0.000)	(0.001)	(0.001)
same college	0.000	0.000	0.000	0.001	0.001
	(0.000)	(0.000)	(0.000)	(0.001)	(0.001)
same grad school	0.024***	0.024***	0.023***	0.027***	0.027***

Includes user-level fixed effect. Standard errors clustered at the user-level. Weighted so that each user's weights sum to 1.

Legend: * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$ Continued on Next Page...

Table 1.26 – Continued

Variable	**Exten.**	**Inten.**	**Ext./Int.**	**Inten.**	**Ext./Int.**
		Absolute	**Absolute**	**Pct.**	**Pct.**
	Primary Sample	**Primary Sample**	**Primary Sample**	**Alt. Sample**	**Alt. Sample**
	1	**2**	**3**	**4**	**5**
	(0.002)	(0.002)	(0.002)	(0.003)	(0.003)
alter friends (100)	-0.000	-0.001***	-0.001***	-0.001**	-0.001**
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
alter friends $(100)^2$	0.000**	0.000***	0.000***	0.000***	0.000***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
alter friends $(100)^3$	-0.000**	-0.000***	-0.000***	-0.000***	-0.000***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
alter friends $(100)^4$	0.000**	0.000***	0.000***	0.000***	0.000***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
constant	-0.009***	0.001	-0.007***	0.000	-0.009**
	(0.002)	(0.001)	(0.002)	(0.001)	(0.003)
R2	0.122	0.121	0.122	0.133	0.133
N	1,017,089	1,017,089	1,017,089	1,438,699	1,438,699

Includes user-level fixed effect. Standard errors clustered at the user-level. Weighted so that each user's weights sum to 1.

Legend: * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$

Chapter 2

Gun For Hire: Delegated Enforcement and Peer Punishment in Public Goods Provision

2.1 Introduction

The title character of the 1950s television western, Paladin, is described as a "gentleman" and "accomplished warrior" who "insists the *rule of law* be enforced, rejecting man-to-man frontier justice" (Hirschman, 2000). His calling card read simply, "Have Gun. Will Travel." The lawlessness of the "wild west" can be described in modern terms as "peer-to-peer punishment" in which scores are settled between parties, often with inefficient punishment. Paladin encouraged cooperative behavior by providing order through reason and, failing that, force. He was a gun for hire and was portrayed as a costly but superior alternative to shootouts, feuds, and endless retribution.

This paper provides a theoretical model and experimental analysis of Paladin. A special problem faced by small self-governing groups is that punishments meted out by members can often be quite deleterious. Peers often punish to the extent that they erase any gains brought on by the punishment, resulting in short-run net losses. Although a strong demand for peer-to-peer punishments exists in the laboratory setting, we observe little of this type of justice in the real world. Instead we often observe the development of delegated or appointed parties that sanction bad behavior. Consider the following examples: the homeowners' association, the building superintendent, the soccer coach, the department head, committee chair, the parent teacher association, and synagogue or church elders. These authorities are created and often funded by a subset of the people and institutions that they monitor. One suspects that these mechanisms arise because they are more efficient than the alternative of vigilante justice. We see people and institutions choosing a hired gun instead of punishing each other for infractions in the real world, but a thorough investigation of the two different mechanisms has not been conducted.

We propose the "Gun For Hire" mechanism as one example of a third party mechanism based on a simple rule of punishment of noncompliance. The rule is low cost to enforce, in equilibrium results in no punishments and full compliance, and when off the equilibrium path typically results in punishments that will be small. A central feature of the gun for hire is that the enforcer does not need to perfectly document all the noncompliance. The gun for hire only needs to know the exact actions of the two largest deviators from compliance. In many instances, the biggest deviators (think of loudest neighbors, worst teachers, most truant volunteers) are easy to identify. Moreover, the enforcer does not need to punish all non-compliant people, just the single biggest cheater. Finally, the punishment need not be large. It only needs to be just big enough that the most non-compliant person would rather have been the second most non-compliant person. If the second most non-compliant person is best responding to his or her environment, the two most non-compliant choices should be nearly identical, meaning that in expectation this difference should be trivial. It follows that punishments off the equilibrium path will likely be small. Hence, even if our mechanism requires some experience to reach equilibrium, the costs along that road should be minimal.

We show that (as seen on TV) a simple mechanism (a gun for hire) is an efficient and desirable substitute for lawless peer-to-peer punishment. We first use an experiment to show that our gun for hire mechanism works when it has been imposed exogenously on our subjects. Next, we show that subjects will choose to implement the gun for hire mechanism, and that it will work even when only a subset of subjects choose to implement the mechanism. Last, we show that even when vigilante justice is always available when there is a gun for hire, subjects discontinue the use of peer-to-peer punishments. That is, the mere presence of a centralized enforcement mechanism makes people less willing to employ vigilante justice.

In an effort to demonstrate the potential for research in this area, we use a series of linear public goods games to examine whether the gun for hire mechanism works first when it is exogenously assigned and second when it is endogenously chosen. In all of our games, subjects are randomly assigned to a group of four people, in which they are asked to allocate an endowment between a public good and a private good. We experiment with four enforcement regimes that players can use to discourage free riding. Subjects have three types of punishment conditions: an exogenously imposed gun for hire, an endogenously chosen gun for hire, peer-to-peer punishments only, or both peer-to-peer punishment with an endogenously chosen gun for hire. The gun for hire mechanism is meant to be just one of any number of examples of small scale self-policing devices; it is a stylized version of the homeowner's association or building superintendent. By looking at the peer-to-peer and gun for hire separately and jointly, we can identify their relative welfare effects.

We find that when our gun for hire is exogenously imposed it immediately improves welfare. When the mechanism is endogenously chosen, there is significant demand for the gun for hire both when it is the only punishment option and when it is offered alongside peer-to-peer punishment. Welfare, as measured by group net earnings (that is earnings minus the costs of punishment), is significantly improved when groups can choose to hire a gun compared to when they can only peer-to-peer punish. Welfare is also improved when subjects can choose to hire a gun in addition to peer-to-peer punish compared to when they can only peer-to-peer punish. Furthermore, when both types of punishment are available and the gun is hired, the costs of peer-to-peer punishment decline precipitously.

In sum, when peer punishment is the only option, individuals use it, often with negative welfare consequences. When given the option of a centralized punishing mechanism, players prefer this to taking justice into their own hands; they cease to engage in peer punishment, and welfare improves dramatically. While our model and results are highly stylized, we will argue that the experimental observation is suggestive of a common real-world phenomenon: inefficient social institutions (such as peer-to-peer punishment) can be easily supplanted by lower cost, more efficient mechanisms that delegate enforcement.

2.2 Background

In previous experiments on costly peer-to-peer punishment, subjects can pay a fee to reduce the payoff to another subject in their group *only* once. While this type of peer-to-peer punishment leads to higher contributions to the public good, the effects on group welfare (group earnings minus punishment costs) have been ambiguous. Egas and Riedl (2008), Gachter et al. (2008), Herrmann et al. (2008), Botelho et al. (2007), Fehr and Gachter (2002), Fehr and Gachter (2000), and Ostrom et al. (1992) all found decreases in net earnings in the short run, while Masclet et al. (2003) found that adding a single round of punishment increased net earnings.[1] If the peer-to-peer punishment is repeated over many periods (50 periods of play, rather than 10) with the *same* groups intact Gachter et al. (2008) found a welfare improvement. In this case, it is possible that repeated interaction created reputation or reciprocity concerns that may have partially driven this result.

Notice that a single round of costly punishment does not take into account the possibility for revenge. When an opportunity for counter-punishment is added net earnings are dramatically reduced, as found by Denant-Boemont et al. (2007) and Nikiforakis (2008).[2] Hence, multiple rounds of costly punishment can create disastrous revenge cycles (Nikiforakis and Engelmann, 2011). One way to lower the costs is to allow non-monetary punishments such as disapproval messages or exposing only low contributors (Noussair and Tucker, 2005; Masclet et al., 2003; Savikhina and Sheremeta, 2010). Another avenue for lower costs is to allow subjects to threaten sanctions before contribution decisions are made (Masclet et al., 2011; Bochet and Putterman, 2007). The

[1] See (Vesterlund, 2012) for a review. Also note that if punishments are only carried out when at least two members of the group request them, then over time there is a welfare gain (Casari and Luini, 2009). While, if players receive a noisy signal of other group members' behavior then the addition of punishment is not only detrimental to welfare, but also decreases contributions to the public good. See Grechenig et al. (2010).

[2] Cinyabuguma et al. (2006) found that if subjects are not given information about who specifically punished them, then net earnings increase. This restriction on the information basically makes revenge motivated second round punishments impossible.

fact that people enjoy expressing their disapproval is convincingly shown by Fudenberg and Pathak (2010), who demonstrate that subjects still engaged in costly punishment even though it was not observed until the end of 10 rounds of play. In such a case, punishment logically could have no effect during the game. If people enjoy punishing, and if costly punishment is the only tool available, then the negative welfare effects of costly punishment are likely to be exacerbated by revenge cycles. It may be that people enjoy punishing because they believe these punishments carry out justice. If this is the case, a person may not actually want to carry out the punishing herself and would gladly hire a third party to mete out justice on her behalf.

This literature suggests that to improve welfare, we need to curb the enjoyment of punishment and prevent peer-to-peer revenge cycles. When the streets are full of vendettas, and desparados are roaming the frontier looking for a fight for fun, what do the town folks do? They call Paladin. That is to say that a natural method for solving both these problems is "hiring" or "appointing" someone to discipline the group. Note that by delegated we don't necessarily mean someone outside the group, but simply mean a commonly recognized conduit for complaints, who monitors and metes out punishments. The punishments need not be more severe than those available by peer-to-peer punishment. The key is that discipline is centralized and a credible threat. Some previous work has already shown that central coordination of punishment can be welfare improving both theoretically (Kube and Traxler, 2011; Boyd et al., 2010; Sigmund et al., 2010; Steiner, 2007) and in experiments (Baldassarri and Grossman, 2011; Dickinson and Villeval, 2008; Falkinger et al., 2000; Yamagishi, 1986).

Yamagishi's experiment is most closely related to our gun for hire. Yamagishi allowed subjects to first play the public goods game and then contribute to a punishment fund that punished the lowest contributor to the public good. Unlike our mechanism, Yamagishi's punishment size was not related to the size of deviation from compliance. Yamagishi finds public contributions were higher under punishment, but welfare was only improved under certain cost schemes. Although these results lend credence to the idea that there is a welfare gain from a delegated sanctioning mechanism, we believe that choosing the amount of punishment after the public contribution decisions is fundamentally different than choosing to hire a delegated mechanism before the public goods game has taken place. We also see our study as building on Yamagishi's insights by making punishments sensitive to the severity of the infraction.

If delegated punishment is the solution, will people voluntarily submit to a gun for hire? Clearly many positive examples exist in the real world on both a large and small scale, such as the police regulating public safety, the EPA assessing fines for emissions, the PTA socially penalizing those who don't sell raffle tickets, the building superintendent speaking to the noisy neighbors, or the department chair cracking down on bad teaching. There have been some experiments in which subjects have been able to choose if they would like to be punished either by each other (Sutter et al., 2010; Ertan et al., 2009; Gurerk et al., 2006; Botelho et al., 2007; Decker et al., 2003) or by a third party (O'Gorman et al., 2009; Kosfeld et al., 2008; Guillen et al., 2007; Tyran and Feld, 2006). These authors have found that, in some cases, subjects choose to allow punishing. Many of these experiments have made the implementation of a punishing mechanism monetarily costless. Monitoring, however, typically requires some resources or opportunity cost. By contrast, we make our punishment mechanism costly when it is endogenously chosen, but the cost is less than the gain realized through cooperation.

Our intuition for the hired gun comes from two sources. First is simple observation of real life mechanisms. Speeding tickets from police officers are not generally issued to everyone on the freeway going over the speed limit but rather to the fastest car on the road. To avoid a speeding ticket, one only needs to be the second fastest car. That is, enforcement of compliance in the real world often focuses first, and often exclusively, on the most egregious violators. The second source of intuition is from the Keynesian p-beauty contest games

(Ho et al., 1998; Nagel, 1995). Imagine a game in which the winner of a prize is the person who guesses a number between 0 and 100 that is closest to two-thirds of the average of the others' guesses. As long as there is common knowledge of rationality, people will realize that (through iterated deletion of dominated strategies) the only way for everyone to be two-thirds of the average is if they all guess 0, which is the Nash equilibrium. Our mechanism turns this intuition upside down. Here the "loser" will be the largest free rider (that is, the one who gained the most by deviating from full compliance), and the penalty will be enough for her to wish she had been the second biggest cheater rather than the biggest. This gives everyone the incentive to be the second biggest cheater. The only set of choices in which everyone can avoid being the biggest cheater (again with common knowledge of rationality) is full compliance.

2.3 The Games

The experiment contains five different public goods games. We use the the linear public goods game with four players as the basic framework for each game, so we will begin by explaining the rules for this game.

2.3.1 The Linear Public Goods (LPG) game

Subjects are given an "automatic payment" of \$1 (to reduce within experiment income effects, as will be seen later) and an endowment of 5 tokens that they allocate between a public good and a private good. Each token invested in the public good pays a return of \$2 to all group members for an aggregate social return of \$8. Each token invested in the private good pays a return of \$3 to only the individual who made the investment. Let g_i be player i's contribution to the public good. The earnings for a subject for a period are:

$$\pi_i = 1 + 3(5 - g_i) + 2\sum_{j=1}^{4} g_j$$

A selfish profit-maximizing player would choose to set $g_i = 0$ and if all players are selfish they will each earn \$16. The group welfare maximizing level of contribution is $g_i = 5$. If all players choose this amount, their earnings would be \$41 each. After all subjects have chosen g_i they are given anonymous information about the contribution to the public good, private good, and initial LPG earnings for each of their group members. This game will act as the basic framework for our other games. Next, we will explain how our gun for hire mechanism works when it has been exogenously imposed on players. We will call this the "Gun Hired" game.

2.3.2 The Gun Hired (GH) game

Subjects are given an "automatic payment" of \$0.50 (to reduce within experiment income effects, as will be seen later) and an endowment of 5 tokens that they allocate between a public good and a private good. Subjects are informed that they are playing with a third party punishment mechanism that will punish the lowest contributor to the public good.

What The Hired Gun Shoots

The gun for hire mechanism simply takes a deduction from the lowest contributor to the public good. The size of the bullet fired by the hired gun varies with the size of the infraction from the group behavior. The size of the deduction is set so as to make the lowest contributor to the public good just slightly worse off (in terms of net subgame payoff) than the second lowest contributor to the public good. In our mechanism the two payoffs will differ by the value of one unit of the private good, \$3.

Formally, let g_z denote the contribution of the lowest contributor to the public good, $g_z = min\{g_1, g_2, g_3, g_4\}$. If there is a tie for the lowest contributor, then all those who tied will be punished. Let g_y denote the second lowest contribution to the public good, $g_y = min\{g_1, g_2, g_3, g_4 \backslash g_z\}$. The size of the punishment will be the difference between the initial payoffs of player z and player y plus a constant, M. We set M equal to the cost from taking one token of the player's private good, so $M = \$3$.

The punishment for player z is equal to:

$$P = \pi_z - \pi_y + 3 = 3(g_y - g_z) + 3$$

In the special case in which all the players choose the same level of contribution to the public good, but still give below full contribution ($g_i = g_j < 5\ \forall i, j$), all the subjects are punished P_0. We set P_0 to \$3, the payoff from contributing a token to the private good. Lastly if all 4 subjects contribute the full 5 tokens to the public good, then no one is punished. To summarize, when the gun is hired, the size of the shot fired is equal to:

$$P = \begin{cases} 3 & \text{if} \quad g_i = g_j < 5 \text{ for all } i, j \\ 0 & \text{if} \quad g_i = 5 \text{ for all } i \\ 3(g_y - g_z) + 3 & \text{for lowest contributor(s) in other cases} \end{cases}$$

Subjects are aware of this punishment mechanism when they make their choices of contribution to the public good, g_i. After all players have chosen g_i, they are given anonymous information about the contribution to the public good, private good, initial LPG earnings, size of punishment (if any), and final net payoffs for each of their group members.

GH Equilibrium

Notice that any choice of $g_z < g_y$ will result in the subject earning \$3 less than player y. This choice is strictly dominated by a choice of $g_i = g_y + \epsilon > g_y$, where $\epsilon > 0$ is the smallest positive increment of g. The choice of $g_i = g_y + \epsilon > g_y$ will result in no punishment. That is, the best response of the lowest contributor is to change g_z to be just slightly higher than g_y. If all subjects reason this way, it is never a best response to set $g_i = 0$. Knowing that all subjects will not choose to set g_i to zero, a subject will choose g_i equal to the next discrete amount above zero, $g_i = 1$. But then knowing that everyone else is using similar reasoning, subjects will want to choose the next discrete amount above $g_i = 1$, and so they need to move to $g_i = 2$. In short, the best response for any player is to find what the lowest level of contribution is, and to set their contribution slightly above it. The only fixed point is full contribution to the public good $g_i = 5$. See the online Appendix for the generalized model and proofs.

The game is like a p-beauty contest (Nagel, 1995) in reverse. Each player is trying to guess the lowest amount given by the others in her group and then wants to give the closest contribution above that amount possible. This thought process eventually pushes all the players to contribute all of their endowment to the public good. Each player should choose $g_i = 5$ and will earn \$40 in the game (this does not include the \$0.50 fixed payment). We will compare earnings under this exogenously imposed mechanism to earnings when the mechanism is endogenously chosen. In the next section we explain how the game is played when we allow players to pay a fee to implement the gun for hire.

2.3.3 The Gun For Hire (G4H) game

The Gun For Hire (G4H) game is very similar to the the Gun Hired (GH) game. The only difference is that we add a pre-play stage 1. In stage 1, each subject is given an endowment of 4 tokens worth \$0.25 each. Subjects choose e_i, $0 \leq e_i \leq 4$, to contribute to the "hiring fund." If the sum of the 4 person group's contributions reach a threshold of 8 tokens, a delegated punishment mechanism will be implemented in stage 2. Subjects' stage 1 earnings equal the number of tokens they kept multiplied by \$0.25. Overpayments for hiring the gun are not refunded to the subjects. If the threshold for hiring is not met, subjects are refunded their e_i and earn \$1 in stage 1. Thus, if the gun is hired, then subjects play the aforementioned Gun Hired game. Note that it costs 8 tokens per group or 2 tokens per person (on average) to hire the gun. Because each token is worth \$0.25, the average cost is \$0.50 per person, which is equivalent to the automatic payment in the Gun Hired (GH) game. Also if the gun is not hired, then all tokens offered in stage 1 are refunded to each player, which is equivalent to the value of the "automatic payment" in the basic LPG game.

In short, the Gun For Hire game goes as follows in stage 1: subjects choose whether to hire the third party punishing mechanism. If they do not hire, they play the regular linear public goods game (LPG) with equilibrium earnings in the sub-game of \$15. If they do hire, they play the gun hired (GH) game with sub-game equilibrium earnings of \$40. A subject should be willing to pay any amount less than or equal to the gain from hiring the gun (\$25) to hire the gun. We have set the total group cost of hiring the gun to only \$2 per group. Any combination of contributions summing to exactly \$2 will be an equilibrium of the stage 1 game (Bagnoli and Lipman, 1989; Bagnoli and McKee, 1991; Marks and Croson, 1998).

Any two players could pay for the punishment mechanism, so one could interpret the implementation of our mechanism as requiring 50% of the group to agree on implementation. The average cost of the gun per person should be \$0.50, and with the gun hired second stage earnings which should be \$40, the resulting average earnings should be \$40.50 per subject in the G4H game.

2.3.4 The Peer-to-Peer (P2P) game

Our peer-to-peer punishment game is similar to that of previous experiments (see Fehr and Gachter, 2002; Cinyabuguma et al., 2006; Herrmann et al., 2008; Gachter et al., 2008). Subjects first play the LPG game with an automatic payment of \$1 (again to reduce income effects), then are given anonymous information about the contribution to the public good, private good, and about initial LPG earnings for each of their group members. At this point, each player i can pay \$1 to assign a punishment point to another player j, which we write as p_{ij}. Each point assigned reduces player j's payoff by \$3.[3] Final payoff are given by the following expression:

$$\pi_i = 1 + 3(5 - g_i) + 2\sum_{j=1}^{4} g_j - \sum_{j \neq i} p_{ij} - 3\sum_{k \neq i} p_{ki}$$

Given that groups are randomly and anonymously rematched each period, own-profit maximizing subjects should choose to assign zero punishment points to all players ($p_{ij} = 0$), and the game should be the same as the LPG game. The predicted outcome under own-profit maximizing behavior is $g_i = 0$ for all subjects and final earnings per subject of \$16.

It is important to note that the own-profit maximizing equilibria predictions of the P2P and LPG games are the same, but that many previous works have found that subjects behave very differently in these two games. The fact that players engage in punishment at all is surprising, not only because it is not the equilibrium action, but more so because we do not observe much peer-to-peer punishment in many real world situations as found by Balafoutas and Nikiforakis (2011). One reason we observe such high amounts of peer punishment in the lab may be that players were never offered another alternative, such as hiring a delegated punishing mechanism in addition to peer punishments. Our final game allows the use of both a delegated punishment mechanism and peer-to-peer punishments.

[3]The punishment to cost ratio of 3:1 has been employed by many of the previous experiments (e.g. Fehr and Gachter, 2002; Gachter et al., 2008; Herrmann et al., 2008), while some others have employed a 4:1 ratio (Cinyabuguma et al., 2006). For a discussion of the constant ratio versus other punishment regimes see Casari (2005). Previous work has found that a cost to punishment ratio of no lower than 1:3 is necessary to raise public contributions and welfare (Nikiforakis and Normann, 2008; Egas and Riedl, 2008). There is the possibility of earning a negative payoff in the P2P game. Subjects were warned about the possibility of negative payoffs in the instructions and were told that they would never owe money at the end of the experiment; and that at minimum they would be paid \$7. In only 3 cases did a subject earn a negative amount in a period.

2.3.5 The Gun For Hire and Peer-to-Peer (G4H/P2P) game

The last game combines the G4H and P2P games. In stage 1, subjects are given 4 tokens, and they make contributions toward a hiring fund. If the sum of those contributions is greater than 8 tokens, then a gun is hired and subjects get \$0.25 for each token they kept. If the gun is not hired, stage 1 earnings are \$1. Subjects are informed of their stage 1 earnings, group contributions to the hiring fund, and whether the gun has been hired. In stage 2, subjects get 5 tokens to contribute to either a public or private good. If the gun was hired, then the lowest contributor(s) to the public good will be punished by the delegated punishment mechanism. In stage 3, subjects are given anonymous details of group members' contributions to the public good, the private good, and their initial earnings (earnings before punishments from the hired gun mechanism). They also learn the size of punishment from the mechanism (if any) and the net earnings for each subject in their group. At this point, subjects can choose to assign peer-to-peer punishments to their group members. Again, subject i chooses an amount of punishment points to assign to player j. Each point player i assigns costs player i \$1, and reduces the payoff of player j by \$3.

Again, own-profit maximizing subjects would assign zero punishment points, leading to predictions identical to the G4H game: subjects hire the gun in the first stage, and fully contribute in the second stage. Average per person earnings would be \$40.50 per person. Table 2.1 summarizes the equilibrium predictions for each of these games. We see that theoretically the gun for hire mechanism whether chosen or imposed (GH, G4H, G4H/P2P) should result in better provision of the public good, and higher average earnings than the linear public goods game (LPG) or the peer-to-peer (P2P) game.

Table 2.1: Equilibrium Predictions

Game	Public Contribution	Punishment Points	Total Net Earnings*
LPG: Linear Public Goods	0	na	$16.00
GH: Gun Hired	5	na	$40.50
G4H :Gun For Hire	5	0	$40.50
P2P: Peer-to-Peer	0	0	$16.00
G4H/P2P: Gun For Hire and Peer-to-Peer	5	0	$40.50

*Total net earnings are earnings minus costs of punishment and plus automatic payments.

2.4 Procedures

There are two equally valid views of what is the “natural” baseline. The first is that the the LPG game is the baseline and the P2P is an intervention. The second takes vigilante justice as an ever present option, and so the baseline should be a game with peer-to-peer (P2P) punishments available. We conduct two sets of experiments using both the LPG, and the P2P games as baselines.

Each session involved 12 subjects and 20 periods: 10 periods of a baseline game (either LPG or P2P) followed by 10 periods of a game with punishment (either GH, G4H, P2P, or P2P/G4H). Each treatment is a set of two games, and there are a total of 5 treatments: (1) LPG-GH, (2) LPG-G4H, (3) LPG-P2P, (4) P2P-P2P, and (5) P2P-G4H/P2P.[4] Each treatment was conducted at least 3 times for a total of 36 subjects per treatment. The LPG-GH treatment was conducted 4 times for a total of 48 subjects. We have a total of 192 subjects. Each session was conducted using z-tree software (Fischbacher, 2007), lasted under 90 minutes and subjects earned $28 on average.

To minimize repeated game effects, participants were randomly and anonymously re-matched into a new group of 4 participants at the beginning

[4]One may be curious why we did not run all permutations of combinations of these games. The reason is we were primarily interested in starting with a baseline world (either LPG or P2P) and then adding on an additional punishment option. So for example if we had run LPG-G4H/P2P then we would have started in a baseline world with no punishment and then added two punishment options. We find this is an interesting manipulation but it was not the focus of this paper.

of each period.[5] Subjects were given the instructions for the first 10 periods of play, a quiz, and then played that game for 10 periods. All participants had to correctly answer the quiz questions before moving on. This is done again for the last 10 periods. To remove experimenter effects, all sessions were run by the same person. Subjects could earn up to \$46 in each period, so they were informed that they would be paid for a single randomly selected period from the 20 periods in the session.[6]

The instructions were written in neutral language by referring to the public good as the "BLUE investment", the private good as the "RED investment", the delegated punishment mechanism as "the computer simulated administrator", and referring to all punishments as "deductions." Full instructions and screen shots are available from the authors in the online Appendix.[7]

[5]The use of a random strangers matching protocol should minimize the effect of reputations and contagion because subjects do not know who they are playing with nor if they have played with them before or will play with them in the future. However it is possible that a player may play against the same subject or even within the same group multiple times, and it is also possible for a player to be affected through contagion. We believe these reputation and contagion effects did not have a significant effect on our results, and as a robustness check we have asked whether our results hold when looking at only the first period of play following our baseline games (e.g. Period 11 of the (1) LPG-GH, (2) LPG-G4H, (3) LPG-P2P, (4) P2P-P2P, and (5) P2P-G4H/P2P). The results are of the same sign and generally remain statistically significant in all cases.

[6]To choose the random period after the end of the 20th period, a subject was given a 20 sided die. The subject was asked to verify the die had 20 sides, and then to roll and announce the outcome on the die out loud.

[7]We included a number of examples in the text of the instructions and in the tests of understanding that we made each subject pass before moving on to actual game play. As pointed out by one of our very helpful reviewers we explicitly mention an example of full contribution to the public good in the GH, G4H and G4H/P2P instructions while we neglect to use this same example in the LPG and P2P instruction sets. This was an unfortunate and unintended oversight, however we do not believe it had any impact on our results. In particular, we have data from another experiment on the same subject pool where subjects played a linear public goods game and and the instructions explicitly mention the full contribution example. The text of the instructions read "Example 3: Imagine you invested your 5 tokens this way: 0 in the RED and 5 to the BLUE investment. Also imagine the other group members invest 5, 5, and 5 to the BLUE investment." We found that average public contribution in these LPG games was 1.5 tokens for the first 10 periods for the 80 subjects who had this written in their instructions. In contrast our subjects who had LPG instructions without this example gave an average of 1.7 tokens for the first 10 periods. The difference between these two means is not statistically significant at standard significance levels (if anything it appears the example pushed contributions down). This

2.5 Results

The "natural" baseline for our experiments is either a world without any punishment options, the LPG game, or a world with only vigilante justice, the P2P game. We will begin by exploring the results for a world without any punishments available in the first 10 periods.

2.5.1 Baseline: LPG

When we begin in a world with no punishment options we are first interested in testing if our gun for hire mechanism will work when it is exogenously imposed in the GH game. After we have shown that the mechanism works we will show that it is still effective when it is endogenously hired in the G4H game. Last we will show that when we compare the endogenously chosen third party mechanism (G4H) to the endogenously chosen vigilante justice (P2P) we find that welfare is greatly improved under the third party mechanism.

Exogenously Imposed Mechanism: GH

We first use a within subjects comparison to test whether our mechanism can fix the free riding that has built up in the first 10 periods of the LPG game. Looking at Figure 2.1 and Table 2.2 we can see that in period 1 to 10 subjects contribute an average of 1.56 tokens per period and that there is a trend toward more free-riding as they repeat the game. When subjects have the gun for hire imposed on them in the last 10 periods of the LPG-GH treatment there is an immediate jump in contributions and the average contribution to the public good rises to 4.57 tokens per period. Earnings increase from \$23.82 per period in the LPG game to \$37.28 per period in the GH game.

leads us to believe that the lack of inclusion of this example in the instructions had no effect. Nonetheless, we regret not having been more consistent in our choice of examples in the instructions.

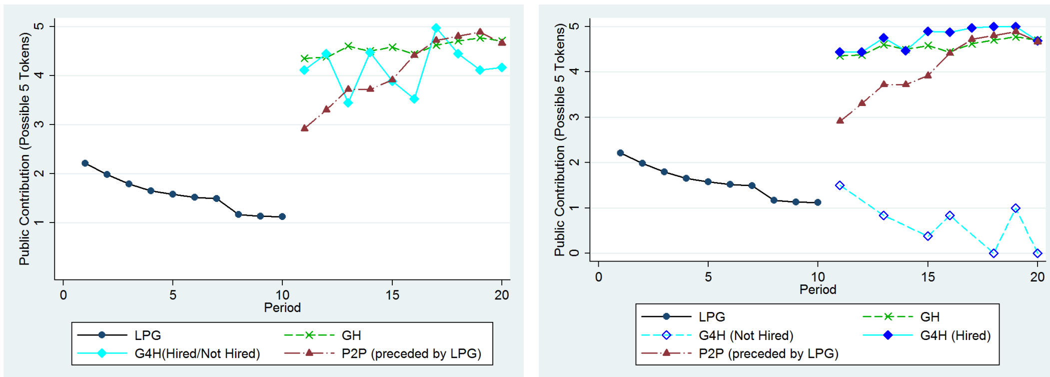
This jump is immediate, as can be seen in Figure 2.2. Clearly the mechanism has had the desired effect of reducing free-riding and increasing earnings even after punishments have been taken away.

Next we can use between subject comparisons to see if our mechanism performs well both when it is imposed (GH) and when subjects have to pay a cost to implement it (G4H). Although the equilibrium of the G4H game is to implement the gun for hire, subjects may not always immediately realize this fact. For subjects to hire the gun, they must believe that the cost of implementing the delegated punishment mechanism will be outweighed by the gains from reduced free-riding. Subjects appear to believe this − they hire the mechanism 85% of the time in last 10 periods of LPG-G4H. In fact our gun for hire is over-paid for. There are multiple equilibria for the hiring stage, such that any combination of contribution to the hiring fund that total exactly \$2 (8 tokens) is a Nash equilibrium. Yet, we only observe the groups paying exactly \$2 a mere 10% of the time in the G4H game, 90% of the time the gun for hire is over-paid for.

We can show that the mechanism improves public contributions both when it is imposed and when it is hired. In the left hand panel of Figure 2.1 we can compare average contributions to the pubic good in the G4H game (note this includes both when the gun has been hired and when it has not) to the GH game. The contributions are slightly higher, but insignificantly, in the GH game.[8] In the right panel of the Figure 2.1 we can see public contributions in the G4H game divided into when the gun was hired versus when it was not hired. Here we see that when subjects successfully hire the gun they actually contribute more on average than when it was imposed on

[8]The difference is not significant using a Kolmogorov-Smirnov test at the session level when comparing public contributions in the GH to all the public contributions in the G4H game (all meaning both when the gun is hired and not hired), $p = 0.237$. We use a Kolomogrov-Smirnov test because we only have 4 observations at the session level for the GH and 3 at the session level for the G4H games. Our results differ from those of Sutter et al. (2010) who found higher contributions for endogenously chosen mechanisms.

them, but this difference is not statistically significant.[9] Clearly, when the gun is not successfully hired, which is only 15% of the time, the subjects contribute much less.[10]



Notes: In this Figure we show the average per person contribution to the public good by treatment out of a possible 5 tokens for treatments which began with 10 periods of the Linear Public Goods (LPG) game, a game with no punishment mechanism. In the left hand panel the "Gun For Hire" (G4H) treatment is the average over both when the hired gun mechanism was and was not implemented. In the right hand panel the "Gun For Hire" (G4H) treatment is divided into when the hired gun mechanism was implemented (solid line) and was not implemented (dashed line).

Figure 2.1: Contributions to the Public Good after LPG

Next we ask whether the gun for hire mechanism also improves net earnings. Net earnings are earnings after the costs of hiring and punishment have been deducted. We can see in Figure 2.2 and Table 2.2 that earnings are \$37.28 on average in the GH treatment while they are \$35.44 in the G4H treatment. These differences in earnings are not statistically significantly different, so it appears that the opportunity for the mechanism alone raises earnings.[11] If we divide the G4H game into when the gun was hired or not hired, we find

[9] Kolmogorov-Smirnov test at the session level when the gun has been hired $p = 0.265$.

[10] Kolmogorov-Smirnov test at the session level when the gun has ***not*** been hired in the G4H versus the GH game, $p = 0.047$

[11] For earnings $p = 0.237$ using a Kolomogrov Smirnov test at the session level. We use a Kolomogrov-Smirnov test because we are only have 3 observations at the session level for the P2P and G4H games.

Table 2.2: Average Earnings per Subject with baseline LPG

Game (Periods)	**Earnings (Dollars)**	**Public Contribution (5 tokens)**	**P2P Costs (All)**	**G4H Costs (All)**
LPG (1-10)	23.82	1.56	na	na
GH (11-20)	37.28	4.57	na	1.06
G4H (11-20)	35.44	4.16	na	1.35
Hired (85%)	38.12	4.74	na	1.57
Not Hired (15%)	19.55	0.71	na	na
P2P (11-20)	30.69	4.11	5.86	na

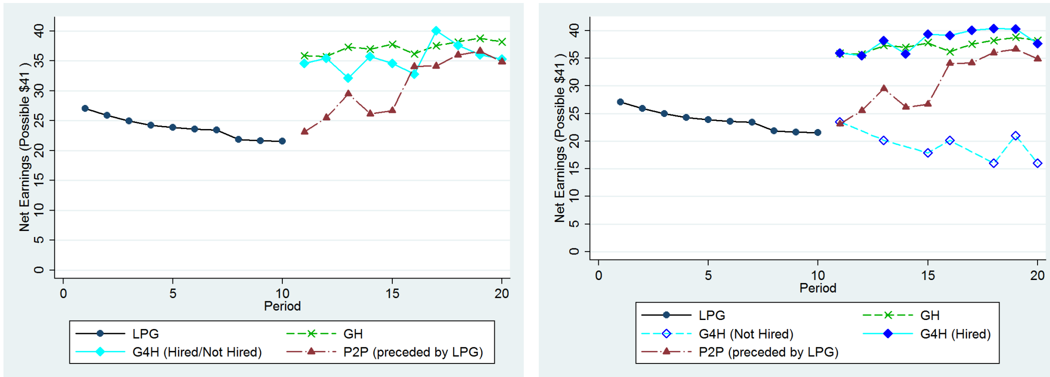
Note: 10 Periods of each game per Session, 3 Groups per Session, 4 Subjects per Group

that when subjects hire the gun their earnings are around \$38.12 per person, while if they do not hire, they earn an average of \$19.55.

Result 1*: Subjects are willing to pay a cost to submit to a delegated punishment mechanism. In the G4H game the delegated punishment mechanism is implemented 85% of the time, and groups over-pay for this implementation in most cases. Welfare, as measured by average individual net earnings, is similar when the gun has been endogenously hired in the G4H game versus when it has been exogenously imposed in the GH game.*

Endogenously Chosen Punishments: G4H versus P2P

We have shown that the gun for hire works when it has been either exogenously assigned or when it has been endogenously paid for. We next compare the effectiveness of G4H to P2P. Looking at Table 2.2, we see that average per-person earnings in Periods 11 to 20 are higher in the LPG-G4H treatment (\$35.44 overall: \$38.12 when gun is hired and \$19.55 when not hired), than in the LPG-P2P treatment (\$30.69 in the last 10 periods). Table 2.3 provides an overview of how average earnings are shaped in each treatment of this experiment. The variable G4H takes the value 1 when subjects are playing the



Notes:In this Figure we show the average per person earnings after punishment deductions by treatment out of a possible $41 for treatments which began with 10 periods of the Linear Public Goods (LPG) game, a game with no punishment mechanism. In the left hand panel the "Gun For Hire" (G4H) treatment is the average over both when the hired gun mechanism was and was not implemented. In the right hand panel the "Gun For Hire" (G4H) treatment is divided into when the hired gun mechanism was implemented (solid line) and was not implemented (dashed line).

Figure 2.2: Average Per Subject Net Earnings after LPG

G4H game and zero when they are playing the P2P game after periods 1-10 of LPG. Playing the G4H game instead of the P2P game raises earnings by $4.76 per period on average including when the gun was not hired.[12]

There are two possible reasons for the increased average earnings: increase average giving and decreased average punishment costs. Table 2.2 shows

[12]The same patterns of significance can be shown in Kolomogrov Smirnov test at the session level. For our regression to properly identify effects we must make two assumptions. We must assume that the session is a random variable, which it should be given random assignment of treatments to sessions. We can also show that it passes a Breusch and Pagan Lagrangian multiplier test for random effects. Second, we must also assume no correlation between the session and the observable right hand side variables which in our case are the period, and the treatment dummies. Again this assumption should be met given random assignment of treatment to sessions. We conduct our analysis at the session level (6 sessions per regression: 3 G4H game and 3 P2P game in Regression 1; 3 G4H/P2P game and 3 P2P game in Regression 2) because we have used a strangers matching protocol which means that individuals can play with the same group members multiple times during a session. If we cluster at the session level we are assuming that actions are independent across sessions, which seems a safe assumption given sessions never have the same persons and treatments are assigned randomly. Estimating equations and results with standard errors clustered at the individual level have smaller standard errors and are available in the online appendix.

that average giving was nearly identical in P2P (4.11) and G4H (4.16). However, as Figure 1 2.1 shows, this average masks a great deal of heterogeneity across treatments. When the delegated punishment mechanism is hired, average giving is higher in G4H.[13] As a result, punishment in G4H are small ($1.35 per subject), especially in comparison to P2P ($5.86 per subject). Thus, lower punishments are primarily responsible for the increased efficiency.

Table 2.3: Determinants of Earnings in Rounds 11-20 by treatment

	After LPG	**After P2P**
G4H	4.76***	
	(1.14)	
G4H/P2P		8.98**
		(3.07)
Period	0.90***	0.49***
	(0.10)	(0.09)
Constant	16.77***	15.25***
	(1.76)	(2.59)
N	720	720
Wald Chi-Squared	97.14***	38.03***

Notes: G4H= 1 if subject in the G4H condition in rounds 11-20, and G4H= 0 if subject in the P2P condition in rounds 11-20 after playing LPG in 1-10. Similarly, G4H/P2P= 1 if subject in the G4H/P2P condition in rounds 11-20, and G4H/P2P= 0 if subject in the P2P condition in rounds 11-20. Linear random effects models. Clustered standard errors in parentheses. Standard errors clustered by session. *** $p < .01$, ** $p < .05$, * $p < .10$ significance

The gain in earnings between the two treatments is illustrated graphically in Figure 2.2. In the left panel are earnings both when the gun is hired and when it is not hired averaged together. These earnings are almost always higher than earnings with P2P punishments. In the right panel, we see that

[13]Comparing session level public contributions in the P2P to when the gun is hired in G4H the public contributions in G4H conditional on hiring are statistically significantly higher than those in P2P using a Kolomogrov Smirnov test $p = 0.09$.

when the gun is hired, average per subject earnings are always higher than those under P2P punishments.

Result 2: *Welfare, as measured by average individual net earnings, is higher in the G4H treatment than the P2P treatment. When the mechanism is hired, the use of the delegated punishing mechanism both improves public contributions, and lowers costs as compared to allowing peer-to-peer punishments.*

One can see in both Figures 1 and 2 that the advantage of G4H over P2P diminishes with time, that is earnings and public contributions in the G4H treatment and P2P treatment appear to converge in the last 5 periods of play. In the next section we will see if this is also the case when we start off in a baseline world with P2P punishments.

2.5.2 Starting From Vigilante Justice: Welfare in Period 11-20 following P2P in Periods 1-10

It has been argued that peer-to-peer punishment "plays an important role in real life" (Fehr and Gachter, 2000). If such peer-to-peer punishment is indeed natural and often occurring then we should use the P2P game as our baseline rather than the setting without any punishment opportunities. One may wonder if subjects will still hire the third party mechanism when they know that they can use vigilante justice. We found that subjects hired the gun in the G4H/P2P game 72% of the time. Similar to the LPG baseline, 80% of the time the gun is over-paid for in the G4H/P2P game.

Table 2.4 shows that average per person earnings in Periods 11 to 20 are higher in the G4H/P2P treatment ($31.87 overall: $36.14 when gun is hired and $20.77 when not hired), than they are in the P2P treatment ($22.89).[14]

[14]One might notice that earnings in the last 10 periods of P2P-P2P are $22.89 which is much lower than earnings in the last 10 periods of LPG-P2P which were ($30.69). This is

Average net earnings are 40% higher in periods 11-20 in the P2P-G4H/P2P treatment than they are in the P2P-P2P treatment, a significant increase.[15]

Additionally, in the regression reported in the second column of Table 2.3 one can see that the coefficient on the treatment dummy variable G4H/P2P is positive and significant. Playing the G4H/P2P game instead of the P2P game raises earnings by $8.98 per period on average.[16] Figure 2.3 shows that contributions to the public good in the last 10 periods of the P2P-G4H/P2P treatment seem to rise over time, while the public contributions stay relatively flat in the P2P-P2P treatment. In the left panel of Figure 2.4 we see that averaging over when the gun is hired and not hired subjects have higher per person net earnings in the G4H/P2P treatment than in the P2P treatment in every period. This result is even more clear when one looks to the right hand panel of Figure 2.3 where contributions have been decomposed into when the gun was hired and when it was not hired. Additionally Figure 2.4 shows that earnings trend upwards for the G4H/P2P treatment, while they stay relatively flat in the P2P treatment. Result 3 summarizes.

Result 3: *Subjects are willing to pay a cost to submit to a delegated punishment mechanism even when they know they will have the ability to peer-to-peer punish. In the G4H/P2P game the delegated punishment mechanism is implemented 72% of the time, and groups over-pay for this implementation in most cases. Welfare, as measured by average individual net earnings, is significantly higher in the last 10 periods of the P2P-G4H/P2P treatment than the P2P-P2P treatment.*

---

a surprising difference, especially given the theoretical prediction is that in both games the average earnings should be $16. We would like to thank an anonymous referee for pointing this out, and we see this as a fruitful question for future research.

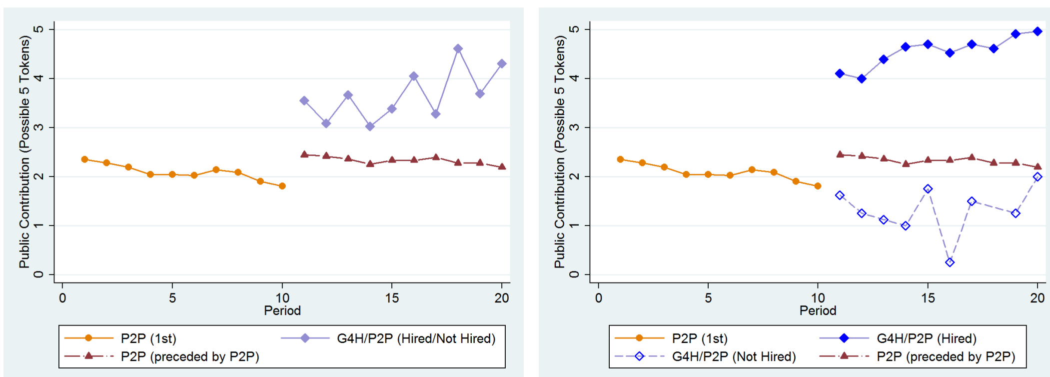
[15]These are statistically significantly different from each other using a Kolomogrov Smirnov test ($p = 0.090$) at the session level.

[16]The variable G4H/P2P takes the value 1 when subjects are playing the G4H/P2P game, and zero when they are playing the P2P game after periods 1-10 of P2P. The same patterns of significance can be shown in Kolomogrov Smirnov test at the session level.

Table 2.4: Average Earnings per Subject after P2P in Periods 1-10

Game (Periods)	Net Earnings (Dollars)	Public Contribution (5 tokens)	P2P Costs (All)	G4H Costs (All)	Total Costs
P2P (1-10)	22.36	2.09	4.07	na	4.07
P2P (11-20)	22.89	2.33	4.74	na	4.74
G4H/P2P (11-20): All	31.87	3.67	1.03	1.43	2.45
Hired (72%)	36.14	4.55	0.63	1.97	2.60
Not Hired (28%)	20.77	1.37	2.08	na	2.08

Note: 10 Periods of each game per Session, 3 Sessions, 3 Groups, 4 Subjects per Group



Notes: In this Figure we show the average per person contribution to the public good by treatment out of a possible 5 tokens for treatments which began with 10 periods of the Peer-to-Peer (P2P) game, a game with a punishment mechanism. In the left hand panel the "Gun For Hire/Peer-to-Peer" (G4H/P2P) treatment is the average over both when the hired gun mechanism was and was not implemented. In the right hand panel the "Gun For Hire/Peer-to-Peer" (G4H/P2P) treatment is divided into when the hired gun mechanism was implemented (solid line) and was not implemented (dashed line).

Figure 2.3: Contributions to the Public Good after P2P

2.5.3 Does Delegated Enforcement Crowd Out Peer Punishment?

We have shown that our G4H mechanism is both implementable and welfare improving when compared to a P2P punishment regime. Next, we show

Notes: In this Figure we show the average per person earnings after punishment deductions by treatment out of a possible $41 for treatments which began with 10 periods of the Peer-to-Peer (P2P) game, a game with a punishment mechanism. In the left hand panel the "Gun For Hire/Peer-to-Peer" (G4H/P2P) treatment is the average over both when the hired gun mechanism was and was not implemented. In the right hand panel the "Gun For Hire/Peer-to-Peer" (G4H/P2P) treatment is divided into when the hired gun mechanism was implemented (solid line) and was not implemented (dashed line).

Figure 2.4: Average Per Subject Net Earnings after P2P

that hiring a gun crowds out the use of peer punishments. If delegated punishment crowds out peer-to-peer punishment, this may be welfare improving. Also, if delegated punishment crowds out peer-to-peer punishment, this will in turn lower any possible motives for for peer-to-peer revenge punishments.[17]

[17]We expect that the use of a delegated punishing mechanism would preclude revenge motives in subsequent rounds of punishment, although we have not allowed for multiple rounds of punishment in our current design. When the mechanism is levying fines, it is not possible for an individual to know who to take revenge on. Imagine if your neighbor was leaving garbage in the common areas of your building. You can either speak with your neighbor directly, or ask the superintendent to speak to your neighbor without mentioning your name. If you speak with your neighbor directly they may take offense, and they may "counter-punish" you by stealing your newspaper. On the other hand if your superintendent speaks with your neighbor, there is no way for your neighbor to know that you commissioned the punishment.

As a helpful anonymous reviewer pointed out subjects could use peer punishment to exact revenge on group members who helped to pay for the gun for hire mechanism during that period. Although we do not reveal who paid for the mechanism, subjects could punish those they suspected of paying to hire. To test for this in our data we looked at subjects who were punished by the hired gun and who also assigned punishment points. There are only 4 instances of this in all our data, so using the peer punishment for revenge on those who helped to hire the mechanism does not appear to be a widely used strategy.

Looking back at the Table 2.4, we can compare the behavior of subjects who could only peer punish in periods 11-20 (P2P) to those who were also allowed to hire a gun and peer punish (G4H/P2P). When we make this comparison we see that peer punishment costs fall from an average \$4.74 in the P2P game to \$1.03 in the G4H/P2P game (\$0.63 when the gun is hired, and \$2.08 when the gun is not hired). The average use of peer punishment when it is the only option is over four times higher than when peer punishment is available alongside the option for a hired gun (\$1.03 versus \$4.74). Like previous work, it is interesting that subjects choose to punish at all because this is not the Nash equilibrium of the one shot game. Recall the hypothesis suggested earlier that subjects enjoy punishing, even when it is costly to them, because they experience utility from enforcing justice. If presented with the option to implement a "just" mechanism, people may prefer this mechanism to vigilante justice. Our subjects may believe that the gun for hire is such a "just" mechanism, and so refrain from punishment when it is offered as an option. Unfortunately this paper is unable to test if subjects have this taste for justice, but we believe it is an important area for further research.

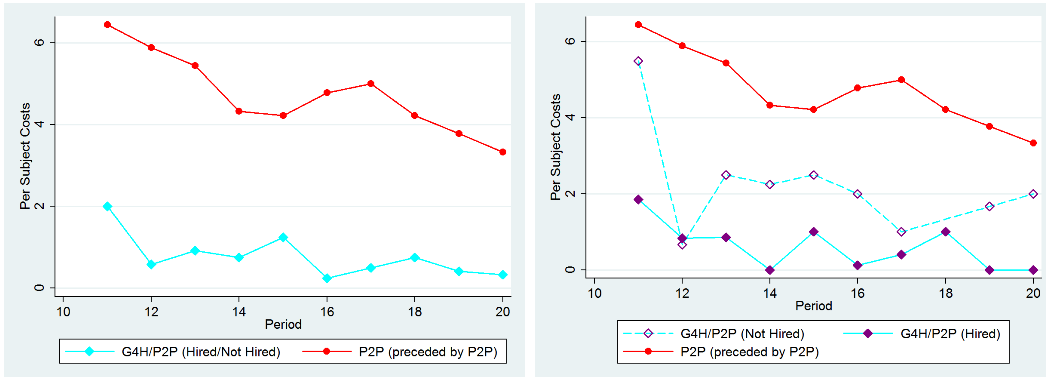
Some may wonder if the gun for hire mechanism is simply a less expensive way to punish than the P2P. In the P2P game, it always costs \$0.33 to punish another player \$1. In the G4H and G4H/P2P games the ratio varies since the punishment depends on size of deviation, and can range as high as \$18. Although it may appear that paying only \$0.50 per person (\$2 in total) to punish \$18 is simply a great value, in our experiments the punishment was usually well below this \$18 size. On average it cost \$0.74 to punish \$1 in the G4H game, and \$0.51 to punish \$1 in the G4H/P2P game, and so punishment was actually cheaper under the P2P mechanism.

Comparing the use of peer punishment by hiring decision in the G4H/P2P treatment alone, we see the average costs of peer punishment fall about 70% when the gun is hired. Figure 2.5 shows the average costs per subject of peer punishments in the P2P game, plus the costs by hiring decision in the G4H/P2P game. In both the P2P and G4H/P2P games, the use of peer punishment is trending downward over time. In the P2P game the costs of punishment are always higher than in the G4H/P2P game, whether a gun has been hired or not. One would expect the use of peer punishments to fall when a gun has been hired, but it is especially surprising that the use of peer punishments is lower even when the gun is not successfully hired. Although there is selection into the hiring of the gun, even when averaging over all groups (those that hired and that did not hire), it appears that merely giving the option for the gun for hire, even when that option is not exercised, decreases peer-to-peer punishments. Furthermore, it is noteworthy that in 4 of the 10 periods (after 10 periods of P2P) when the gun is hired peer punishment costs are equal to zero.

***Result 4**: Use of peer punishments is over four times higher when the option for a delegated mechanism is not available. When the delegated mechanism is implemented, peer punishment converged to zero by period 19. Delegated punishment crowds out peer-to-peer punishment, resulting in an overall welfare gain.*

2.6 Concluding Remarks

Much of the previous work on punishment in public goods games has concentrated on asking whether groups can govern themselves through the use of peer-to-peer punishments. This line of inquiry does not allow individuals to collectively agree to concentrate the punishment in a recognized authority. In this paper, we show that subjects willingly pay to delegate punishments in a linear public goods game. We offer a stylized version of delegated punishment in our gun for hire mechanism. The mechanism has the properties that only the largest free rider is punished, the size of the punishment is related to the degree of defection from the other group members' behavior, in equilibrium



Notes: In this Figure we show the average per person punishment costs by treatment for treatments which began with 10 periods of the Peer-to-Peer (P2P) game, a game with a punishment mechanism. In the left hand panel the "Gun For Hire/Peer-to-Peer" (G4H/P2P) treatment is the average over both when the hired gun mechanism was and was not implemented. In the right hand panel the "Gun For Hire/Peer-to-Peer" (G4H/P2P) treatment is divided into when the hired gun mechanism was implemented (solid line) and was not implemented (dashed line). The reason there is no data point for Period 18 on the non-implemented (dashed) line in the right panel is that the gun was hired by all groups during Period 18 in all sessions.

Figure 2.5: Average Per Subject Peer Punishment Costs after P2P in Periods 1-10

the mechanism is efficient in the sub-game, and the mechanism is relatively low cost.

When given the opportunity to hire a delegated punishment mechanism, we see the mechanism being implemented over 70% of the time in both the linear public good and peer-to-peer baseline worlds. The likely reason that subjects are so willing to submit to a costly outside authority is that they expect monetary gains from reduced free-riding. These expectations are well-founded, as can be seen by the 15% and 40% increase in welfare when comparing a peer-to-peer punishment regime to those with the option of a gun for hire regime. When subjects can only use peer punishments (P2P) the peer punishment costs are over four times those with a delegated mechanism

(G4H/P2P). Last and most important, we find that when both punishment methods are available (G4H/P2P), subjects lower their use of peer punishments by 70%. The existence of a delegated punishing mechanism crowds out the use of peer punishments.

To our knowledge, this paper is the first to allow subjects to choose between hiring a costly punishment mechanism and using peer-to-peer punishments. We have shown that players want to hire the delegated mechanism and that the gun for hire mechanism provides a low cost solution to the problem of free-riding. Interestingly, although the delegated mechanism is itself a public good, it does not appear to suffer from the same level of free-riding as observed in the subsequent LPG game. The reason may be that the cost of hiring is fairly low as compared to the potential gains in payoffs. This is analogous to the way we pay taxes or fees to fund delegated punishing mechanisms in general. Often these fees and penalties are small, as in our mechanism. Further research exploring how players react to changes in the cost of implementing the mechanism could be illuminating. Although formally our gun for hire was an external third party, it is clearly an important and desirable next step for research to investigate a more general set of ways individuals can delegate authority. For instance, the recognized authority can be internal to the group, and the enforcers's conformity with enforcement rules a choice variable. This would be most interesting, of course, in the default domain of peer-to-peer punishment. In our experiment the third party always executes punishments exactly as dictated by the mechanism, but if the gun for hire was an actual person one may worry about abuse of power, making it critical to keep the power of the authority relatively weak, as cogently pointed out by Binmore (2005). Work exploring how a human third party authority might abuse power is also an important further research question. The ultimate research goal suggested by our study is to understand how easily small self-governed groups can innovate ways to avoid the inefficiencies of peer-to-peer punishment.

This paper illustrates that under reasonable conditions individuals prefer to pay to be governed by a delegated punishment mechanism rather than use peer-to-peer punishments. The gun for hire mechanism is just one example of a low-cost device that can deter free-riding behavior in a public goods game, improve welfare, and crowd out the use of deleterious peer-to-peer punishments.

In short, when Paladin comes to town, vigilante justice is driven out. Have gun. Will travel.

2.7 Acknowledgement

Chapter 2, in full, is a reprint of the material as it appears in the Journal of Public Economics, vol. 96, no. 11-12, December 2012, p. 1036-1046. Gee, Laura; Andreoni, James., 2012. The dissertation/thesis author was the primary investigator and author of this paper.

Chapter 3

The Nature of Giving Time to Your Child's School

3.1 Introduction

It is a fact that households contribute time (via volunteering or fundraising) to their children's schools, but what motivates these actions? Time contributions to the school have both public and private benefits. The public benefit of time contribution is an increase in overall school quality that benefits all the enrolled students. The private benefits may be that parents enjoy spending time with their child, want to act as a role model, gain influence over the direction of school resources, or avoid the guilt of refusing to give their time. Private benefits are not necessarily selfish or egoistic, they are simply benefits which accrue to the individual household rather than to the whole school. When considering why households choose to give their limited time to their children's schools it is important to know if they are motivated by both the public and private benefits for this action. There are many factors that will affect a household's choice to contribute time to the school including but not limited to: feelings of efficacy, social norms, school outreach, the saliency of the school and whether the school has been chosen or is simply the default. Although these are all interesting and important determinants of school time contribution, the focus of this paper is to test whether households care about the private benefits of their time contribution. One way to test this hypothesis is to compare households with multiple children in a single school, to households with multiple children in different schools. A household gains the private benefits of time contribution for each of their own children at a school. Raising the number of own children at a school should raise the marginal private benefits from an hour of time contribution. Using this simplified framework I can test if a a change in the number of own children at the school (holding all else constant) has a statistically significant relationship with time contribution.

Using a nationally representative sample of over 2,500 multi-child households, I find that time contributions move in parallel with the number of children from a single family enrolled in the same school. I control for child, household and school level observable variables, but I am unable to control for unobservable variables (e.g. the salience of the school to the household). I find that households with two children enrolled in the same school have a higher propensity to contribute time than households with a single child in each school. Households with two children in the same school have a propensity to volunteer at the school which is 13 percentage points higher, a propensity to volunteer in the classroom which is 8 percentage points higher, a propensity to fundraise which is 10 percentage points higher, and on average spend 6 more hours per year in these activities than households with a single child in each school. This finding implies that households view time contribution to their child's school as producing a private benefit, and that they take account of this private benefit when making time contribution decisions.

3.2 School Contribution: Private and Public Benefits

One model of how parents choose how much time to give to their children's schools might assume that parents are maximizing household utility subject to a time budget constraint (they make similar decisions about monetary contributions, but that is not the topic of this paper). Each individual household benefits from the total time given by other parents to the school in activities such as chaperoning field trips, fundraising for a new theater, or coaching the softball team. Suppose at a particular school the provision of all these services takes a total of 100 hours of parental contribution. If the parents care only about the final public goods being provided, then they would be happy if other families gave all the 100 hours to the school, and they personally gave 0 hours. Traditional models of public goods predict that a household with these types of preferences will give an inefficiently low amount of time to the school, and that there will be under provision of the public goods. However, empirically we do not observe high levels of free riding behavior by households in the school, nor for public goods in general. Specifically in the sample discussed in this paper 48% of the households volunteered at their child's school, 25% volunteered in their child's classroom specifically, 65% engaged in fundraising, and on average households spent 13.6 hours per year in these activities.

An alternative to these traditional public goods models assumes the household receives a private benefit from the act of giving time to the school. There are many possible private benefits to giving time to a child's school. The private benefit may be a "warm glow" of increasing school quality overall or on a per student basis (Andreoni, 1990, 2007). Alternatively, it may be that the parent enjoys acting as a role model for their child by being a socially responsible, philanthropic community member (Mustillo and Lynch, 2004).[1] Also it could be that the household is able to build relationships at the school that allow them to get preferential treatment, and/or avoid stigma from not giving time. Under any of these interpretations, the time contribution is motivated by preferences over these private benefits, as well as preferences over the public benefit of improved school quality. There is a rich literature about incorporating these types of impure motivations for giving into the public goods model (Kotchen, 2006; Duncan, 2004; Cornes and Sandler, 1994). In these models a consumer chooses to allocate some scarce resource (time or money) between a private good, which benefits only the consumer, and a public good, which gives the consumer some private return while also providing a benefit to the whole community.

It is simple to see how time contribution to the school is a likely candidate for this type of model. Each household chooses between time spent doing other activities (sleeping, paid work, watching television) and time spent contributing to their children's schools (via volunteering or fundraising). Time

[1]The National PTA suggests that a way to help one's child succeed is to "Be a role model; be active in community service yourself or together with your child," which would include giving time to a child's school. (http://www.pta.org/100Ways_brochure-en.pdf)

contributed to the school benefits the individual family through enhancing their child's school experience, role model effects, warm glow, guilt avoidance, and/or through the ability to better allocate school resources to their children. The time contribution also produces a positive externality that improves school quality for all the students. The household contributes the number of hours to the school where the marginal benefit of an hour given to the school is equal to the marginal benefit of an hour spent doing something else, and so that the number of hours contributed is below the total hours in the period of interest (e.g. 8,750 hours in a year). If the marginal benefit of time spent at the school rises, then the household will adjust their choice of hours until the marginal benefits of all activities are once again equal.

Consider a mother who has two children in different schools. When she volunteers at each school she enhances the quality of each school, and gets some extra private benefit from spending time at her children's two schools. Now consider a father who also has two children, but his children are enrolled in the same school. When he volunteers at the single school, he still gets the benefit that each child's school is improved in quality, and the additional private benefit from spending time at his children's single school. It seems likely that the time the father spends volunteering at a single school (where he has two children enrolled) has a higher marginal return than the time spent by the mother with children in two different schools. If households with two children in the same school get a higher return on contribution to the public good, then these households will behave differently than other households and may be more likely to contribute to the public good.[2] If the household is considering

[2]This model of behavior does not take account of the idea that a household with two children in the same school may simply think more about that school, than a household with a single child in a school. This issue of the salience of the school may be driving increased contribution at the school rather than an actual higher private benefit from having two children at the school. This salience may come from being asked to volunteer more often by the school. I attempt to address this particular type of salience in the section titled "The Importance of Being Asked". It is also possible that the salience comes from something other than being asked, and I am unable to include these other measures of salience in this analysis.

the private benefits then their contribution choice should be affected by number of own children at the school.

3.3 Method

3.3.1 Data

The data come from the 2003 National Household Education Survey (NHES) of Parent and Family Involvement (PFI). The NHES PFI is a phone-based survey that was conducted in 1996, 2003, and 2007. The 2003 data are used because the sampling procedure in that wave has multiple child level observations within a household. The survey asks an adult household member questions about the school age children in the household. If there is a single school age child, then that child is a single observation. If there are many school age children, two of those children are chosen randomly to represent two observations. This means that at most there are two observations per household. The hypothesis is that households with multiple children in multiple schools will be less likely to contribute time than a household with multiple children in a single school. Households with an only child are likely to be very different than households with multiple children. Also single child families automatically must enroll only one child in a school, and so they are excluded from this analysis. To test the hypothesis I compare households with two or more children enrolled in the same school to households with two or more children in multiple schools.[3] In 2003 there were 12,426 school-age child observations representing 8,467 households in the original data set. I restrict the sample to households that have two parents with two or more non-homeschooled children. The final data set includes 5,750 child level observations over 2,875

[3]In some cases parents may not enroll two children in the same school because of age differences, for example one child is in kindergarten and the other is in high school. In other cases the family may elect to send children who could be in the same school to different schools, for example one student enrolls in an arts school and the other in a technical school.

households.[4]

3.3.2 Dependent Variables

The dependent variables of interest are (1) whether a household member has volunteered at the child's school (Volunteer), (2) whether a household member has volunteered in the child's classroom (Classroom), (3) whether a household member has engaged in fundraising for the school (Fundraise), and (4) the number of hours spent volunteering or fundraising by the household per year per school (Hours).[5] The first three independent variables are dummy variables, taking the value 1 if the household said they did participate in the activity and 0 if they did not. A household is only asked if they volunteered in the classroom (Classroom) if they answered in the affirmative to volunteering at the school (Volunteer). The fourth dependent variable is the number of hours that the household reported contributing to the specific school (Hours), so it can take any non-negative value.[6]

---

[4] 4,508 child level observations were dropped because there was only a single observation within the household. 64 households were dropped because one or more of their children were not currently enrolled in school, even though they were of school age. 162 households were dropped because one or more of their children were home schooled. 1,725 households were dropped because they did not have both a mother and a father at home (of these 65 households were dropped because the type of parents were coded differently across surveys in the same household). 2,310 households were dropped because they did not have 2 parents and siblings (of these 41 households were dropped because the family makeup differed across surveys in the same household).

[5] The specific questions are "Since the beginning of this school year, have you (or (CHILD)'s (mother/stepmother/foster mother/father/stepfather/foster father/grandmother/grandfather/aunt/uncle/cousin) (or (the) other adult(s) in your household))
d. Acted as a volunteer at the school or served on a committee?
e. Served as a volunteer in (CHILD'S) classroom?
f. Participated in fundraising for the school?"
and
"Since the beginning of this school year, how many hours have you or (CHILD)'s (mother/stepmother/foster mother/father/stepfather/foster father/grandmother/grandfather/aunt/uncle/cousin) (or (the) other adult(s) in your household)) participated in (volunteering) (and) (fundraising) at (CHILD)'s school?"

[6] The reported hours range from 0 to 600. Only 19 households report spending more than 200 hours per year volunteering/fundraising at their child's school. Running the same

3.3.3 Independent Variables

The independent variables were chosen because they have been found in previous studies to be important predictors of charitable giving (time or money). Previous studies have found statistically significant relationships between contribution and many of the variables that were available from the data, these include: race, age, education, employment status, household income, number of children, whether a person lives in a metropolitan area (Freeman, 1997), immigrant status (Osili and Xie, 2009), religiosity (Brown and Ferris, 2007), school enrollment, school racial make-up, school level of free-lunch, school level of English proficiency, and school grade level (Brunner and Sonstelie, 2003). In addition to these variables, 10 other variables that seemed intuitively related to stating that a household contributed to the school were also added to the model. These extra variables were: child's age, child's sex, whether the school is religious or year-round, census region, if the interview was conducted in English, and if it was conducted with the child's mother.[7] The independent variable of greatest interest is whether two children attend the same school. If the household gets a higher return to contributing time through having multiple children in the school, then this will lead to an increase in propensity to contribute time. Thus, in the regression with the contribution variable as the dependent variable, one expects the coefficient on the having two children at the same school to be positive and significant.

A limitation of the NHES PFI is that each household in the sample

analysis excluding those households reporting over 200 hours of contribution does not have a strong effect on the results.

[7]The child's age and sex seem relevant because a parent may adjust their volunteering choice if they feel a child is more vulnerable in the classroom due to their age or sex. A school that is year round may have more volunteering opportunities. A school that is religious may have a stronger school emphasis on volunteering. If the interview was not conducted in English there is a higher chance that the questions may have been misunderstood, so I want to control for this possibility. It seems likely that the mother of the child would have the best records of volunteering activity at the child's school, and so I control for when other people may have responded with less accuracy than the mother. The census region was added because different portions of the country may have different feelings about volunteering activity.

has a maximum of two child level observations, even when there are three or more school age children in the home. The sampling procedure for households with three or more children was to randomly select two of the children for the study. The random nature of this selection method should ensure that the final results still hold in spite of missing data for some children in the household. If there is any effect it should bias the regression coefficients toward zero.[8]

The child level characteristics are age, sex, race (white/black/asian/hispanic), language, and birth place. The school level characteristics are grade level (elementary/middle/high school), religious/non-religious, year-round, private/public, and enrollment (as estimated by the respondent). The household level characteristics include the language and family role of the respondent to the interview, religiosity (as proxied by attendance at a religious event in the past month), whether the child has received free or reduced price lunch, number of siblings, household income, mother's and father's age, education (completed high school/college), and employment status (out of labor force/part-time/full-time). Geographic characteristics are the census region, poverty level (by whether more than 5% of households with children in the area are below the poverty line), and whether the household is located in an urban/suburban/rural area.

3.3.4 Data Analysis

To test how strongly households are influenced by the private benefits of time contribution, I use an ordinary least squares (OLS) regression model with robust standard errors clustered at the household level (since there are

[8]Restricting the sample to only those households with exactly two children (2,858 child level observations or 1,429 household observations) does not change the significance or magnitude of the majority of the results. Households with two children enrolled in the same school have a 13 percentage points higher propensity to volunteer (same as in the larger sample), 10 percentage points higher propensity for volunteering in classroom (vs. 8 percentage points in the full sample), 9 percentage points higher propensity for fundraising (vs. 10 percentage points in the larger sample) and spend 6 more hours in these activities on average (same as in the full sample). The full regression results are included in the online Appendix.

two child level observations per household). I am interested in the average reaction (instead of, for example, the elasticity of time substitution), and so I use the simple linear regression model instead of a Probit or a Tobit. Three of the dependent variables take either a 0 or a 1 for a value (whether the household volunteers, volunteers in classroom, or fundraises), and so these would be candidates for a Probit model. Additionally, the hours reported is censored at 0, so it may appear that a Tobit would be the best choice of model. However, I am interested in the average probability and magnitude of contribution and so I ignore the special nature of the variables by running a simple linear regression instead of a Probit, or a Tobit. The analysis of the linear regression model is less complex, and so those results are presented here. The marginal effects results using Probit and Tobit models are generally of the same sign and significance (with similar magnitudes) for most of the independent variables.[9]

3.3.5 Results

Determinants of Contribution

Table 3.1 reports results from the linear regression model; for brevity only selected results are reported (the full table is in the online Appendix). To test whether contribution to the school is affected by the influence of private benefits, I analyze the coefficient on number of own children enrolled at the school ("2 children who attend same school"). In the sample the average

[9]The marginal coefficient for the effect of having two children in the same school is a rise of 16 percentage points in propensity to volunteer (vs. 13 percentage points in OLS model), an 8 percentage point rise in probability of classroom volunteering (the same as the OLS model), a 12 percentage point rise in the probability of fundraising (vs. 10 percentage points in OLS model), and on average those with two children in the same school spend 10 more hours in these activities (vs. 6 hours in the OLS model). The reason the number of hours is so different, is that the OLS model takes account of all the reporting of 0 hours of volunteering, while the Tobit specifically corrects for censoring at 0 and thus has a much higher coefficient on hours. In the Probit and the Tobit the coefficients on the contribution variables are still significant at the .1% level. These results are included in the online Appendix.

propensity to volunteer was 48%, to volunteer in the classroom was 25%, to fundraise was 65% and on average households spent 13.6 hours per year per school in these activities. Having two children enrolled at the same school has a significant positive effect on all four measures of contribution; raising the propensity to volunteer by 13 percentage points, volunteer in the classroom by 8 percentage points, fundraise by 10 percentage points, and raising the average hours spent in these activities by 6 hours. This implies that the household cares about the private benefit from contribution, which is rising in the number of own of children enrolled at the school.

Table 3.1: Dependent Variables Measures of Time Contribution (standard errors), abridged results

Description	Volunteer	Classroom	Fundraise	Hours
2 children who attend same school	0.13***	0.08***	0.10***	6.34***
	(0.02)	(0.01)	(0.02)	(1.27)
Child born in the US or US territory	0.06*	0.00	0.07*	-0.23
	(0.03)	(0.02)	(0.03)	(1.70)
Child is in elementary school	0.10**	0.18***	0.11***	8.47**
	(0.03)	(0.03)	(0.03)	(2.86)
Child is in middle school	-0.02	0.02	0.05*	1.25
	(0.02)	(0.02)	(0.02)	(1.54)
Child is in public school	-0.12**	0.01	0.02	-0.42
	(0.04)	(0.04)	(0.04)	(3.22)
Child is in school with religious affiliation	0.03	0.05	0.16***	2.86
	(0.04)	(0.04)	(0.05)	(3.40)
Estimated number of students enrolled is 300 to 599	0.02	0.01	0.03	-1.14
	(0.02)	(0.02)	(0.02)	(1.68)
Estimated number of students enrolled is 600 to 999	0.00	-0.02	0.02	1.68
	(0.02)	(0.02)	(0.02)	(1.94)
Estimated number of students enrolled is over 1000	-0.03	-0.04	0.00	-0.93
	(0.03)	(0.02)	(0.03)	(1.97)
Mother has completed high school	0.07*	0.06**	0.02	3.14*
	(0.03)	(0.02)	(0.03)	(1.41)
Mother has completed college	0.06**	0.02	0.04*	0.81
	(0.02)	(0.02)	(0.02)	(1.37)
Mother Employed 35+ hours per week	-0.11***	-0.12***	0.06**	-7.73***
	(0.02)	(0.01)	(0.02)	(1.43)
Mother Employed part time	-0.00	-0.03	0.04	-2.89

legend: * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$ Table Continued on Next Page...

Table 3.1 – Continued

Description	Volunteer	Classroom	Fundraise	Hours
	(0.02)	(0.02)	(0.02)	(1.91)
Father has completed high school	0.10***	0.02	0.03	2.31
	(0.03)	(0.02)	(0.03)	(1.50)
Father has completed college	0.08***	0.06***	0.01	2.71*
	(0.02)	(0.02)	(0.02)	(1.35)
Father Employed 35+ hours per week	0.03	0.01	0.07*	1.96
	(0.03)	(0.02)	(0.03)	(1.15)
Father Employed part time	-0.01	-0.01	-0.01	0.50
	(0.05)	(0.04)	(0.05)	(2.22)
Parent and child have attended a religious event in past month	0.09***	0.05***	0.04**	2.30*
	(0.01)	(0.01)	(0.01)	(1.11)
Child has received free or reduced price lunch	-0.06**	-0.10***	-0.02	-0.77
	(0.02)	(0.02)	(0.03)	(1.81)
Total number of siblings	-0.01	-0.01	-0.01	-0.21
	(0.01)	(0.01)	(0.01)	(0.79)
Household income between 5,000 and 20,000	-0.01	-0.07	-0.04	0.54
	(0.10)	(0.10)	(0.10)	(3.52)
Household income between 20,001 and 50,000	-0.00	-0.08	0.02	2.89
	(0.10)	(0.10)	(0.10)	(3.59)
Household income between 50,001 and 100,000	0.03	-0.09	0.07	3.24
	(0.10)	(0.10)	(0.10)	(3.72)
Household income above 100,000	0.08	-0.06	0.06	6.83
	(0.11)	(0.10)	(0.10)	(4.03)
Constant	0.10	0.13	0.04	-18.71
	(0.15)	(0.13)	(0.15)	(9.63)
R-squared	0.23	0.26	0.16	0.07

legend: * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$ Table Continued on Next Page...

Table 3.1 – Continued

Description	**Volunteer**	**Classroom**	**Fundraise**	**Hours**
Child Level Observations	5750	5750	5750	5750

Note: Results for child, parental, school and geographic level independent variables excluded from table.

legend: * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$

Generally, the results from the analysis of the NHES PFI are in line with previous studies of contribution. In agreement with other studies, contribution (of time and/or money) is positively related to the education level of the mother and the father (Brunner and Sonstelie, 2003; Freeman, 1997; Andreoni et al., 2003), negatively related to full-time employment by the mother (except for fundraising[10]) (Rotolo and Wilson, 2007; Muller, 1995), positively related to religiosity (Brown and Ferris, 2007), and positively related to being a US citizen (Osili and Xie, 2009).

There are some points where the NHES PFI data do not agree with these previous studies. In this data I find that households do not adjust their propensity to contribute in response to changes in enrollment ("estimated number of students enrolled is 300 to 599", "...600 to 999", or "...over 1000"). This is interesting because it may imply that the household does not consider the total public benefit of their time, when making choices about time contribution. If one believes that the public benefit is dependent on total enrollment, then one would expect there to be a statistically significant relationship between the contribution variables and the school enrollment. One reason we do not observe this result may be measurement error in this data set. The NHES PFI reports enrollment as estimated by the respondent, which could be a very noisy measure of actual enrollment. It is possible that there is too much noise in the measurement of this variable to get a statistically significant relationship. This finding of non-significance of enrollment is in contrast to what Brunner and Sonstelie (2003) found when looking at monetary contri-

[10]Interestingly, households with a fulltime employed mother actually have a propensity to contribute via fundraising which is 6 percentage points higher than those with non-fulltime working mothers. The rise in fundraising is puzzling, but there are two possible explanations. The definition of fundraising may be an issue because the survey does not specify if fundraising takes the form of time contribution (e.g. sitting outside the grocery store selling cookies) or money (e.g. buying raffle tickets from the school). The interpretation of the term "fundraising" may cause confusion in the respondents. Another possible explanation is that households with working mothers may have more social connections they can exploit for fundraising (e.g. selling raffle tickets to co-workers), but that having a full time employed mother puts too much of a time constraint on other volunteering activities.

butions to California public schools. Using data from nonprofit contributions (which are only reported if these contributions are above 25 thousand dollars) to schools from organizations like the PTA, Brunner and Sonstelie found that monetary contributions were higher for schools with higher enrollments. Although the authors believe these monetary contributions come primarily from parents, the contributions may also come from local businesses or foundations. The inability to distinguish what exact portion of this giving comes from families with students at the school may be the reason for the divergence between their results and the ones found here. The censoring of contributions below 25 thousand dollars may also explain the difference.

Another point of disagreement is that while numerous previous studies have found a positive and significant relationship between income and contribution (Brunner and Sonstelie, 2003; Feldman, 2010; Freeman, 1997; Andreoni et al., 2003; Hoover-Dempsey et al., 1987); I find no such relationship.[11] Finally, another common theme in previous studies is the finding that contributions rise in the number of children in the household; again I find no such relationship (the coefficient on "Total Number of Siblings" is insignificant for all dependent variables).[12]

Although in general the results from the NHES PFI match previous studies of contribution, the important new distinction to draw is that giving

[11] I do find a negative relationship between child receipt of free or reduced price lunch and volunteering. Households that received free/reduced price lunch were 6 percentage points less likely to volunteer, and 10 percentage points less likely to volunteer in the classroom on average. One may believe that receipt of free/reduced price lunch catches all the effect of having a lower income household, but removing the lunch indicator and adding finer gradations of income still gives no significant relationship between income and contribution (results in online Appendix).

[12] Initially I thought that these studies might be catching the effect of households having a higher propensity to volunteer when they have multiple children enrolled in the same school. To check I ran the model excluding whether the household has two children in the same school, but the coefficient on total number of siblings is still insignificant for three of the four contribution measures. The coefficient on number of siblings for propensity to volunteer becomes significant at the 5% level, but is negative. This sample of two parent, multiple child households with school age children, may not act in the same manner as households in general. Also since this sample looks specifically at multi-child households it does not catch the gains from moving from zero to one child, or from one to two children.

time to the school depends on number of own students enrolled. This implies that when choosing how to allocate time, the household considers the private benefit from the act of giving time to the school.

Private Benefits

The crux of the argument of this paper is that households receive some private benefit from contributing time to the school. These benefits can range from the happiness a parent might get from being a good role model, to avoiding guilt for not volunteering. The NCES data do not have measures of most of these benefits. However, one example of such a benefit is getting better class or course placement for the children enrolled, and information on this is included in the NCES data. To test if such a benefit likely exists I look at whether parents who have given time to the school believe they have a say in their child's placement. Overall 68% of the survey respondents report feeling they have a say in their child's placement. Using whether the household believes they have a say as the dependent variable and one of the contribution variables as a new additional independent variable I run four new models to see if contribution predicts feelings of having a say in placement. Families that contributed via volunteering generally (Model 1) and in the classroom (Model 2), as well as those who engaged in fundraising (Model 3) have a higher propensity to feel they have a say in their child's placement. Interestingly hours (Model 4) does not seem to correspond to a feeling of having a say in child's placement. When predicting if a family feels they have a say in class placement, the magnitude of time contributed does not seem to matter, but rather whether there is any contribution to the school at all.[13] Below is a table reporting only the coefficients on the contribution variables predicting whether the household feels they have a say in placement, the full table with all independent variables is available in the online Appendix.

[13]Running the model on only those households that report positive amounts of hours contributed, there is still no significant coefficient on number of hours.

Table 3.2: Household Has A Say in Student's Placement (standard errors)

Variable	Model 1	Model 2	Model 3	Model 4
Household has volunteered at school	0.05**			
	(0.02)			
Household has volunteered in classroom		0.04*		
		(0.02)		
Household has engaged in fundraising			0.05***	
			(0.02)	
Hours volunteered by household per school				0.00
				(0.00)
Constant	0.52***	0.52***	0.52***	0.53***
	(0.14)	(0.15)	(0.14)	(0.15)
R-squared	0.04	0.04	0.04	0.04
Child Level Observations	5750	5750	5750	5750

Note: Results for child, parental, school and geographic level independent variables excluded from table.
legend: * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$

The Importance Of Being Asked

The act of being asked may increase the propensity to contribute, because a family may feel guilt about refusing such a request or because they may have been unaware of the opportunity to contribute before the invitation. Freeman (1997) found that being asked to volunteer increased the probability of volunteering by almost 50 percentage points, and experimental studies confirm the finding that being asked increases giving (Andreoni and Rao, 2007). The exclusion of whether a family has been asked to give time may affect our interpretation of the results.

Beyond the simple fact that being asked to volunteer may increase the propensity to volunteer for all households, it is likely that families with two children in the same school may be more likely to be asked to volunteer by the school. It may be that a household with multiple children in the same school, find that school more salient because they are contacted by that school more often than if they had only one student enrolled. If this is the case the more active behavior of the school may be driving the results instead of the preferences of the households with multiple children in the same school. After, including whether the household has been informed by the school about volunteering opportunities in the model, I find it has a positive and statistically significant (at the .1% level) coefficient for all 4 measures of contributions (17 percentage points for volunteering, 5 percentage points for classroom, 21 percentage points for fundraising and 4 more hours on average). However, the level of significance and magnitude of the coefficients on "2 children attend same school" is almost exactly the same for all four measures of contribution.Around 90% of the households report being made aware of opportunities to volunteer, and the variable is quite clearly endogenous so I have chosen to report the results without its inclusion in the model. Results with whether the household has been made aware of volunteer opportunities are reported in the online Appendix.

3.4 Conclusion

In the current economic environment schools continue to attempt to keep school quality high in spite of budget cuts. Increased contribution from parents of time is a possible way to keep school resources high when government based resources may be scarce.[14] To encourage parental contribution, it is useful to know how households view the decision about contributing time to the school. Using a national sample of 2,875 households with two or more children, I find that having two children enrolled in the same school significantly increases the propensity to volunteer (in general and in the classroom), fundraise, and the hours contributed to the school. That number of own children enrolled corresponds to propensity to contribute implies that the con-

[14]Previous studies looking at test scores have found no strong positive effect of equalized funding (Downes, 1992), nor of parental volunteering (Houtenville and Conway, 2008), but student test scores are not the only measure of school quality. I am not currently aware of any studies that have used other measures of school quality to predict the effects of parental volunteering.

tribution decision is affected by some private benefits, and that those benefits increase when there are two or more children in the same school. In line with previous studies, school time contribution is associated with higher parental education, lower female workforce participation, and religiosity. Unlike previous studies enrollment, household income and number of children (in general not at a single school) are not statistically associated with more school contribution. Parental time contribution to the school should be seen as an activity which gives both a public and private benefit to the household where private benefits depend positively upon the number of own children enrolled in the school. When households make decisions about time giving, they appear to be strongly influenced by the private benefits of this time contribution.

3.5 Acknowledgement

Chapter 3, in full, is a reprint of the material as it appears in Nonprofit and Voluntary Sector Quarterly, vol. 40 no. 3 June 2011, p. 552-565. Gee, Laura, 2011. The dissertation/thesis author was the primary investigator and author of this paper.

Bibliography

Akerlof, George A., "The Market for "Lemons": Quality Uncertainty and the Market Mechanism," *The Quarterly Journal of Economics*, 1970, *84* (3), pp. 488–500.

Andreoni, James, "Impure Altruism and Donations to Public Goods: A Theory of Warm-Glow Giving," *The Economic Journal*, 1990, *100* (401), 464–477.

_ , "Giving gifts to groups: How altruism depends on the number of recipients," *Journal of Public Economics*, 2007, *91* (9), 1731 – 1749. Celebrating the 20th Anniversary of Bergstrom, Blume, and Varian's "On the Private Provision of Public Goods".

_ **and Justin Rao**, "The Power of Asking: How Communication Affects Selfishness, Empathy, and Altruism," *working paper*, 2007.

_ **, Eleanor Brown, and Isaac Rischall**, "Charitable Giving by Married Couples: Who Decides and Why Does it Matter?," *J. Human Resources*, 2003, *XXXVIII* (1), 111–133.

Aral, Sinan and Dylan Walker, "Identifying Influential and Susceptible Members of Social Networks," *Science*, 2012.

_ **and Marshall Van Alstyne**, "The Diversity-Bandwidth Trade-off," *American Journal of Sociology*, 2011, *117* (1), pp. 90–171.

Babcock, Philip, "From Ties to Gains? Evidence on Connectedness and Human Capital Acquisition," *Journal of Human Capital*, 2008, *2* (4), pp. 379–409.

Bagnoli, M. and M. McKee, "Volutary Contribution Games: Efficient Private Provision of Public Goods," *Economic Inquiry*, 1991, *29.*

Bagnoli, Mark and Barton L. Lipman, "Provision of Public Goods: Fully Implementing the Core through Private Contributions," *The Review of Economic Studies*, 1989, *56* (4), pp. 583–601.

Bakshy, Eytan, Itamar Rosenn, Cameron Marlow, and Lada Adamic, "The Role of Social Networks in Information Diffusion," *Working Paper*, 2012.

Balafoutas, Loukas and Nikos Nikiforakis, "Norm Enforcement in the city: A natural field experiment," *Working Paper*, 2011.

Baldassarri, Delia and Guy Grossman, "Centralized sanctioning and legitimate authority promote cooperation in humans," *Proceedings of the National Academy of Sciences*, 2011, *108* (27), 11023–11027.

Bandiera, Oriana, Iwan Barankay, and Imran Rasul, "Social Connections and Incentives in the Workplace: Evidence From Personnel Data," *Econometrica*, 2009, *77* (4), 1047–1094.

Baum, Joel A. C., Bill McEvily, and Tim J. Rowley, "Better with Age? Tie Longevity and the Performance Implications of Bridging and Closure," *Organization Science*, 2012, *23* (2), 529–546.

Bayer, Patrick, Stephen L. Ross, and Giorgio Topa, "Place of Work and Place of Residence: Informal Hiring Networks and Labor Market Outcomes," *Journal of Political Economy*, 2008, *116* (6), pp. 1150–1196.

Beaman, Lori A., "Social Networks and the Dynamics of Labour Market Outcomes: Evidence from Refugees Resettled in the U.S.," *The Review of Economic Studies*, 2012, *79* (1), 128–161.

Beaman, Lori and Jeremy Magruder, "Who Gets the Job Referral? Evidence from a Social Networks Experiment," *American Economic Review*, September 2012, *102* (7), 3574–93.

Bian, Yanjie, "Bringing Strong Ties Back in: Indirect Ties, Network Bridges, and Job Searches in China," *American Sociological Review*, 1997, *62* (3), pp. 366–385.

Binmore, Ken, *Natural Justice*, Oxford University Press, 2005.

Bochet, Olivier and Louis Putterman, "Not Just Babble: Opening the Black Box of Communication in a Voluntary Contribution Experiment," *Working Paper*, 2007.

Boorman, Scott A., "A Combinatiorial Optimization Model for Transmission of Job Information Through Contact Networks," *The Bell Journal of Economics*, 1975, *6* (1), pp. 216–249.

Botelho, A., G.W. Harrison, L.M.C. Pinto, and E.E. Rutstrom, "Social Norms and Social Choice," *Working Paper*, 2007.

Boyd, Robert, Herbet Gintis, and Samuel Bowles, "Coordinated Punishment of Defectors Sustains Cooperation and Can Proliferate When Rare," *Science*, 2010.

Bramoulle, Yann and Gilles Saint-Paul, "Social networks and labor market transitions," *Labour Economics*, 2010, *17* (1), 188 – 195.

Brown, Eleanor and James M. Ferris, "Social Capital and Philanthropy: An Analysis of the Impact of Social Capital on Individual Giving and Volunteering," *Nonprofit and Voluntary Sector Quarterly*, 2007, *36* (1), 85–99.

Brown, Meta, Elizabeth Setren, and Giorgio Topa, "Do Informal Referrals Lead to Better Matches? Evidence from a Firm's Employee Referral System," *Federal Reserve Bank of New York Staff Reports*, 2012.

Brunner, Eric and Jon Sonstelie, "School finance reform and voluntary fiscal federalism," *Journal of Public Economics*, 2003, *87* (9-10), 2157 – 2185.

Burke, Moira and Robert Kraut, "Using Facebook after losing a job: Differential benefits of strong and weak ties," *Working Paper*, 2012.

Burt, Ronald S., "Structural Holes and Good Ideas," *American Journal of Sociology*, 2004, *110* (2), pp. 349–399.

Calvo-Armengol, Antoni and Matthew O. Jackson, "The Effects of Social Networks on Employment and Inequality," *The American Economic Review*, 2004, *94* (3), pp. 426–454.

_ **, Thierry Verdier, and Yves Zenou**, "Strong and weak ties in employment and crime," *Journal of Public Economics*, 2007, *91* (12), 203 – 233.

Cappellari, Lorenzo and Konstantinos Tatsiramos, "Friends Networks and Job Finding Rates," *Working Paper*, 2010.

Casari, Marco, "On the Design of Peer Punishment Experiments," *Experimental Economics*, 2005, *8* (2), 107–115.

— **and Luigi Luini**, "Cooperation under alternative punishment institutions: An experiment," *Journal of Economic Behavior and Organization*, 2009, *71* (2), 273 – 282.

Castilla, Emilio J., "Social Networks and Employee Performance in a Call Center," *American Journal of Sociology*, 2005, *110* (5), pp. 1243–1283.

Cinyabuguma, Matthias, Talbot Page, and Louis Putterman, "Can Second-Order Punishment Deter Perverse Punishment," *Experimental Economics*, 2006, *9* (3).

Cornes, Richard and Todd Sandler, "The comparative static properties of the impure public good model," *Journal of Public Economics*, 1994, *54* (3), 403 – 421.

Cui, Ai-Xiang, Zimo Yang, and Tao Zhou, "Roles of Ties in Spreading," *Working Paper*, 2012.

Decker, Torsten, Andreas Stiehler, and Martin Strobel, "A Comparison of Punishment Rules in Repeated Public Good Games: An Experimental Study," *Journal of Conflict Resolution*, 2003, *47* (6), 751–772.

Denant-Boemont, Laurent, David Masclet, and Charles N. Noussair, "Punishment, Counterpunishment, and Sanction Enforcement in a Social Dilemma Experiment," *Economic Theory*, 2007, *33* (1).

Dickinson, David and Marie-Claire Villeval, "Does monitoring decrease work effort?: The complementarity between agency and crowding-out theories," *Games and Economic Behavior*, 2008, *63* (1), 56 – 76.

Downes, Thomas, "Evaluating the Impact of School Finance Reform on the Provision of Public Education: The California Case," *National Tax Journal*, 1992, *45*, 405–419.

Duncan, Brian, "A theory of impact philanthropy," *Journal of Public Economics*, 2004, *88* (9-10), 2159 – 2180.

Egas, Martijn and Arno Riedl, "The economics of altruistic punishment and the maintenance of cooperation," *Proceedings of the Royal Society B: Biological Sciences*, 2008, *275* (1637), 871–878.

Elliott, James R., "Social Isolation And Labor Market Insulation," *Sociological Quarterly*, 1999, *40* (2), 199–216.

Ertan, Arhan, Talbot Page, and Louis Putterman, "Who to punish? Individual decisions and majority rule in mitigating the free rider problem," *European Economic Review*, 2009, *53* (5), 495 – 511.

Falkinger, Josef, Ernst Fehr, Simon Gachter, and Rudolf Winter-Ebmer, "A Simple Mechanism for the Efficient Provision of Public Goods: Experimental Evidence," *The American Economic Review*, 2000, *90* (1), pp. 247–264.

Fehr, Ernst and Simon Gachter, "Cooperation and Punishment in Public Goods Experiments," *The American Economic Review*, 2000, *90* (4), 980–994.

_ and _ , "Altruistic punishment in humans," *Nature*, 2002.

Feldman, Naomi, "Time is Money: Choosing Between Charitable Activities," *American Economic Journal: Economic Policy*, 2010.

Ferrara, Emilio, Pasquale De Meo, Giacomo Fiumara, and Alessandro Provetti, "The role of strong and weak ties in Facebook: a community structure perspective," *International Journal of Social Network Mining*, 2012, *1* (1), 6790.

Fischbacher, Urs, "z-Tree: Zurich toolbox for ready-made economic experiments," *Experimental Economics*, 2007, *10* (2).

Freeman, Richard B., "Working for Nothing: The Supply of Volunteer Labor," *Journal of Labor Economics*, 1997, *15* (1), S140–S166.

Friedkin, Noah, "A test of structural features of granovetter's strength of weak ties theory," *Social Networks*, 1980, *2* (4), 411 – 422.

Fudenberg, Drew and Parag A. Pathak, "Unobserved punishment supports cooperation," *Journal of Public Economics*, 2010, *94* (1-2), 78 – 86.

Gachter, Simon, Elke Renner, and Martin Sefton, "The Long-Run Benefits of Punishment," *Science*, 2008, *322* (5907), 1510–.

Galenianos, Manolia, "Hiring Through Referrals," *Working Paper*, 2011.

Grabowicz, Przemyslaw A., Jos J. Ramasco, Esteban Moro, Josep M. Pujol, and Victor M. Eguiluz, "Social Features of Online Networks: The Strength of Intermediary Ties in Online Social Media," *PLoS ONE*, 01 2012, *7* (1), e29358.

Granovetter, Mark S., "The Strength of Weak Ties," *American Journal of Sociology*, 1973, *78* (6), pp. 1360–1380.

_ , *Getting a Job: A Study of Contacts and Careers*, 2nd ed., The University of Chicago Press, 1995.

Grechenig, Kristoffel, Andreas Nicklisch, and Christian Thoni, "Punishment Despite Reasonable Doubt: A Public Goods Experiment with Sanctions Under Uncertainty," *Journal of Empirical Legal Studies*, 2010, *7.*

Guillen, Pablo, Christiane Schwieren, and Gianandrea Staffiero, "Why feed the Leviathan?," *Public Choice*, 2007, *130*, 115–128. 10.1007/s11127-006-9075-3.

Gurerk, Ozgur, Bernd Irlenbusch, and Bettina Rockenbach, "The Competitive Advantage of Sanctioning Institutions," *Science*, 2006, *312* (5770), 108–111.

Harrigan, Nicholas, Palakorn Achananuparp, and Ee-Peng Lim, "Influentials, novelty, and social contagion: The viral power of average friends, close communities, and old news," *Social Networks*, 2012, (0), –.

Herrmann, Benedikt, Christian Thoni, and Simon Gachter, "Antisocial Punishment Across Societies," *Science*, 2008, *319* (5868), 1362–1367.

Hirschman, Elizabeth, *Heroes, monsters and messiahs: movies and television shows as the mythology of American culture*, Andrews McMeel Publishing, 2000.

Ho, Teck-Hua, Colin Camerer, and Keith Weigelt, "Iterated Dominance and Iterated Best Response in Experimental "p-Beauty Contests"," *The American Economic Review*, 1998, *88* (4), pp. 947–969.

Holzer, Harry J., "Informal Job Search and Black Youth Unemployment," *The American Economic Review*, 1987, *77* (3), pp. 446–452.

Hoover-Dempsey, Kathleen V., Otto C. Bassler, and Jane S. Brissie, "Parent Involvement: Contributions of Teacher Efficacy, School Socioeconomic Status, and Other School Characteristics," *American Educational Research Journal*, 1987, *24* (3), 417–435.

Houtenville, Andrew J. and Karen Smith Conway, "Parental Effort, School Resources, and Student Achievement.," *Journal of Human Resources*, 2008, *43* (2), 437 – 453.

Ioannides, Yannis M. and Adriaan R. Soetevent, "Wages and Employment in a Random Social Network with Arbitrary Degree Distribution," *The American Economic Review*, 2006, *96* (2), pp. 270–274.

_ **and Linda Datcher Loury**, "Job Information Networks, Neighborhood Effects, and Inequality," *Journal of Economic Literature*, 2004, *42* (4), pp. 1056–1093.

Jackson, Matthew O., "Chapter 12 - An Overview of Social Networks and Economic Applications," in Alberto Bisin Jess Benhabib and Matthew O. Jackson, eds., , Vol. 1 of *Handbook of Social Economics*, North-Holland, 2011, pp. 511 – 585.

_ **, Tomas Rodrguez Barraquer, and Xu Tan**, "Social Capital and Social Quilts: Network Patterns of Favor Exchange," *American Economic Review*, Forthcoming.

Jones, Jason J., Jaime E. Settle, Robert M. Bond, Christopher J. Fariss, Cameron Marlow, and James H. Fowler, "Inferring Tie Strength from Online Directed Behavior," *PLOS ONE*, 2012.

Karlan, Dean, Markus Mobius, Tanya Rosenblat, and Adam Szeidl, "Trust and Social Collateral," *The Quarterly Journal of Economics*, 2009, *124* (3), 1307–1361.

Kezdi, Gabor, "Robust Standard Error Estimation in Fixed-Effects Panel Models," *Hungarian Statistical Review Special*, 2004.

King, Gary and Langche Zeng, "Logistic Regression in Rare Events Data," *Political Analysis*, 2001, *9* (2), 137–163.

Kosfeld, Michael, Akira Okada, and Arno Riedl, "Institution Formation in Public Goods Games," *forthcoming, The American Economic Review*, 2008.

Kotchen, MatthewJ., "Green Markets and Private Provision of Public Goods," *Journal of Political Economy*, 2006, *114* (4), 816–834.

Kube, Sebastian and Christian Traxler, "The Interaction of Legal and Social Norm Enforcement," *Journal of Public Economic Theory*, 2011, *13* (5), 639–660.

Leicht, Kevin T. and Jonathan Marx, "The Consequences of Informal Job Finding for Men and Women," *The Academy of Management Journal*, 1997, *40* (4), pp. 967–987.

Lin, Nan, "Social Networks and Status Attainment," *Annual Review of Sociology*, 1999, *25*, pp. 467–487.

_ **, Paul W. Dayton, and Peter Greenwald**, "Analyzing the Instrumental Use of Relations in the Context of Social Structure," *Sociological Methods and Research*, 1978, *7* (2), 149–166.

Loury, Linda Datcher, "Some Contacts Are More Equal than Others: Informal Networks, Job Tenure, and Wages," *Journal of Labor Economics*, 2006, *24* (2), pp. 299–318.

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

_ **and Steven R. Lerman**, "The Estimation of Choice Probabilities from Choice Based Samples," *Econometrica*, 1977, *45* (8), pp. 1977–1988.

Marks, Melanie and Rachel Croson, "Alternative rebate rules in the provision of a threshold public good: An experimental investigation," *Journal of Public Economics*, 1998, *67* (2), 195 – 220.

Marmaros, David and Bruce Sacerdote, "Peer and social networks in job search," *European Economic Review*, 2002, *46* (45), 870 – 879.

Marsden, Peter V. and Elizabeth H. Gorman, "Social Networks, Job Changes, and Recruitment," in Ivar Berg, Arne L. Kalleberg, Ivar Berg, and Arne L. Kalleberg, eds., *Sourcebook of Labor Markets*, Plenum Studies in Work and Industry, Springer US, 2001.

Masclet, David, Charles N. Noussair, and Marie Claire Villeval, "Threat and Punishment in Public Good Experiments," *Working Paper*, 2011.

_ **, Charles Noussair, Steven Tucker, and Marie-Claire Villeval**, "Monetary and Nonmonetary Punishment in the Voluntary Contributions Mechanism," *The American Economic Review*, 2003, *93* (1), 366–380.

Mayer, Adalbert, "The structure of social networks and labour market success," *Applied Economics Letters*, 2012, *19* (13), 1271–1274.

McEvily, Bill, Jonathan Jaffee, and Marco Tortoriello, "Not All Bridging Ties Are Equal: Network Imprinting and Firm Growth in the Nashville Legal Industry, 19331978," *Organization Science*, 2012, *23* (2), 547–563.

Montgomery, James D., "Social Networks and Labor-Market Outcomes: Toward an Economic Analysis," *The American Economic Review*, 1991, *81* (5), pp. 1408–1418.

_ , "Job Search and Network Composition: Implications of the Strength-Of-Weak-Ties Hypothesis," *American Sociological Review*, 1992, *57* (5), pp. 586–596.

Muller, Chandra, "Maternal Employment, Parent Involvement, and Mathematics Achievement among Adolescents," *Journal of Marriage and the Family*, 1995, *57* (1), 85–100.

Munshi, Kaivan, "Chapter 23 - Labor and Credit Networks in Developing Economies," in Alberto Bisin Jess Benhabib and Matthew O. Jackson, eds., , Vol. 1 of *Handbook of Social Economics*, North-Holland, 2011, pp. 1223 – 1254.

Nagel, Rosemarie, "Unraveling in Guessing Games: An Experimental Study," *The American Economic Review*, 1995, *85* (5), 1313–1326.

Nikiforakis, Nikos, "Punishment and counter-punishment in public good games: Can we really govern ourselves?," *Journal of Public Economics*, 2008, *92* (1-2), 91 – 112.

_ **and Dirk Engelmann**, "Altruistic punishment and the threat of feuds," *Journal of Economic Behavior and Organization*, 2011, *In Press, Corrected Proof*, –.

_ **and Hans-Theo Normann**, "A comparative statics analysis of punishment inpublic good experiments," *Experimental Economics*, 2008, *11*, 358–369. 10.1007/s10683-007-9171-3.

Noussair, Charles and Steven Tucker, "Combing Monetary and Social Sanctions to Promote Cooperation," *Economic Inquiry*, 2005, *43* (3), 649 – 660.

O'Gorman, Rick, Joseph Henrich, and Mark Van Vugt, "Constraining free riding in public goods games: designated solitary punishers can sustain human cooperation," *Proceedings of the Royal Society B: Biological Sciences*, 2009, *276* (1655), 323–329.

Onnela, J.-P., J. Saramki, J. Hyvnen, G. Szab, D. Lazer, K. Kaski, J. Kertsz, and A.-L. Barabsi, "Structure and tie strengths in mobile communication networks," *Proceedings of the National Academy of Sciences*, 2007, *104* (18), 7332–7336.

Osili, Una Okonkwo and Jia Xie, "Do Immigrants and Their Children Free Ride More Than Natives?," *American Economic Review: Papers and Proceedings*, 2009, *99* (2), 28–34.

Ostrom, Elinor, James Walker, and Roy Gardner, "Covenants With and Without a Sword: Self-Governance is Possible," *The American Political Science Review*, 1992, *86* (2), 404–417.

Reagans, Ray and Bill McEvily, "Network Structure and Knowledge Transfer: The Effects of Cohesion and Range," *Administrative Science Quarterly*, 2003, *48* (2), pp. 240–267.

Rider, Christopher I., "Networks, Hiring, and Attainment: Evidence from Law Firm Dissolutions," *SSRN*, 2011.

Rotolo, Thomas and John Wilson, "The Effects of Children and Employment Status on the Volunteer Work of American Women," *Nonprofit and Voluntary Sector Quarterly*, 2007, *36* (3), 487–503.

Sacerdote, Bruce, "Chapter 4 - Peer Effects in Education: How Might They Work, How Big Are They and How Much Do We Know Thus Far?," in Stephen Machin Eric A. Hanushek and Ludger Woessmann, eds., , Vol. 3 of *Handbook of the Economics of Education*, Elsevier, 2011, pp. 249 – 277.

Sarah, John Wilson Mustillo and Scott M. Lynch, "Legacy Volunteering: A Test of Two Theories of Intergenerational Transmission," *Journal of Marriage and Family*, 2004, *66* (2), 530 – 541.

Savikhina, Anya and Roman M. Sheremeta, "Visibility of Contributions and Cost of Information: An Experiment on Public Goods," *Working Paper*, 2010.

Schmutte, Ian M., "Job Referral Networks and the Determination of Earnings in Local Labor Markets," *Working Paper*, 2010.

Shue, Kelly, "Executive Networks and Firm Policies: Evidence from the Random Assignment of MBA Peers," *Chicago Booth Research Paper No. 11-46; Fama-Miller Working Paper*, 2012.

Sigmund, Karl, Hannelore De Silva, Arne Traulsen, and Christoph Hauert, "Social learning promotes insistitutions for governing commons," *Nature*, 2010.

Simon, Curtis J. and John T. Warner, "Matchmaker, Matchmaker: The Effect of Old Boy Networks on Job Match Quality, Earnings, and Tenure," *Journal of Labor Economics*, 1992, *10* (3), pp. 306–330.

Steiner, Jakub, "A Trace of Anger is Enough: on the Enforcement of Social Norms," *Economics Bulletin*, 2007, *8*, 1–4.

Sutter, M., S. Haigner, and M. Kocher, "Choosing the Carrot or the Stick? Endogenous Institutional Choice in Social Dilemma Situations," *Review of Economic Studies*, 2010, *77.*

Tassier, Troy, "Labor Market Implications of Weak Ties," *Southern Economic Journal*, 2006, *72* (3), pp. 704–719.

Topa, Giorgio, "Social Interactions, Local Spillovers and Unemployment," *The Review of Economic Studies*, 2001, *68* (2), 261–295.

_ , "Chapter 22 - Labor Markets and Referrals," in Alberto Bisin Jess Benhabib and Matthew O. Jackson, eds., , Vol. 1 of *Handbook of Social Economics*, North-Holland, 2011, pp. 1193 – 1221.

Tyran, J. and L. Feld, "Achieving Compliance when Legal Sanctions are Non-deterrent," *Scandinavian Journal of Economics*, 2006, *108*, 1467–9442.

Ugander, Johan, Lars Backstrom, Cameron Marlow, and Jon Kleinberg, "Structural diversity in social contagion," *Working Paper*, 2012.

van der Leij, Marco and Sanjeev Goyal, "Strong Ties in a Small World," *Review of Network Economics*, 2011, *10* (1).

Vesterlund, Lise, "Voluntary Giving to Public Goods: Moving Beyond the Linear VCM," *Working Paper*, 2012.

Wei, Thomas E., Victoria Levin, and Lindsay M. Sabik, "A referral is worth a thousand ads: Job search methods and scientist outcomes in the market for postdoctoral scholars," *Science and Public Policy*, 2012, *39* (1), 60–73.

Yakubovich, Valery, "Weak Ties, Information, and Influence: How Workers Find Jobs in a Local Russian Labor Market," *American Sociological Review*, 2005, *70* (3), pp. 408–421.

Yamagishi, Toshio, "The Provision of a Sanctioning System as a Public Good," *Journal of Personality and Social Psychology*, 1986, *51* (1), 110 – 116.

Zenou, Yves, "Explaining the Black/White Employment Gap: The Role of Weak Ties," *Working Paper*, 2011.