THREE ESSAYS IN EMPIRICAL PUBLIC ECONOMICS

by

ADNAN QADIR KHAN

A thesis submitted to the

Department of Economics

in conformity with the requirements for

the degree of Doctor of Philosophy

Queen's University

Kingston, Ontario, Canada

November, 2009

Abstract

This dissertation explores three questions in empirical public economics: we investigate the impact of social networks on labour market outcomes in the first essay; we explore the determinants of volunteering behaviour and estimate the effect of employment on volunteering in the second essay; and we examine the impact of political and fiscal decentralization on public provision in the third essay. In each case, we provide consistent estimates by utilizing an exogenous source of variation in key economic outcomes introduced by randomized policy experiments in the first two essays and by a natural experiment in the third essay. In the first essay, we find that among social networks, weak ties have a significant effect on labour market outcomes but strong ties do not have. In the second essay, we find that employment has a significant effect on volunteering behaviour, and that the effect varies in different contexts and depends on the precise channels through which the two are connected. In the final essay, we find that decentralization has a big effect on public provision. But we also find that decentralization affects different public goods differently, and that the key to its impact lies in the incentives facing politicians at the local level.

Co-Authorship

Chapter 4 of this thesis is part of a joint research project with Prof. Asim Khwaja of Harvard Kennedy School, and Prof. Ali Cheema of Lahore University of Management Sciences.

Dedication

To my wife Samia: For your constant support, encouragement and love. This would not have been possible without you!

To our daughter Rabeal: Born in the first month of my graduate studies, you deserve as much credit as me for sharing the trials and tribulations of graduate life with me, and for inspiring me with your love!

To my parents, Dr. Abdul Qadir and Shakila Qadir, who have taught me to serve the public interest with integrity and dedication!

To my homeland, Pakistan, which deserves peace and prosperity after all that it has gone through!

Acknowledgments

I am greatly indebted to Charles Beach, Steven Lehrer and Robin Boadway for their guidance, support, and insights. While Charles Beach and Steven Lehrer supervised my work, Robin Boadway was a great source of inspiration, encouragement and guidance throughout my four years at QED.

I thank Asim Khwaja, Ali Cheema, Sumon Majumdar, Susumu Imai, Huw Lloyd-Ellis, Jan Zabojnik and James MacKinnon for their constant support and encouragement.

I thank the following seminar participants for providing me useful comments: the Task Force Meeting on Decentralization, Initiative for Policy Dialogue, Columbia University; Queen's Economics Department; and the Canadian Economic Association Conference Meeting.

Finally, I very grateful for to my wife, daughter, and parents for their love and, most of all, their patience.

Table of Contents

Abstract i

Co-Authorship ii

Dedication iii

Acknowledgments iv

Table of Contents v

List of Tables viii

List of Figures xii

Chapter 1:

Introduction 1

1.1 Key Research Questions and Potential Contribution 2

Chapter 2:

Impact of Social Networks on Labour Market Outcomes 6

2.1 Literature, Research Questions and Methodology 6

2.2 Description of Program and Data 12

2.3 Results 31
2.4 Conclusion 43

Chapter 3:

Experimental Evidence on the Effect of Employment on Volunteering 46

3.1 Introduction 46
3.2 Description of Program and Sample Selection 48
3.3 Literature on Volunteering 54
3.4 Model 60
3.5 Empirical Strategy 63
3.6 Estimation Results 71
3.7 Interpretation of Results 79
3.8 Conclusion 82

Chapter 4:

Impact of Decentralization on Provision of Public Services: Evidence from Pakistan 84

4.1 Introduction 84
4.2 Literature 87
4.3 Context of Decentralization in Pakistan 90
4.4 Data and Identification Strategy 97
4.5 Estimation Results 105
4.6 Difference-in-Difference-in-Difference (D-D-D) Estimation 112
4.7 Identification of Political Channels 117

4.8 Conclusion 121

Chapter 5:

Conclusions and Further Work 124

5.1 Policy Implications 125

5.2 Ongoing and Future Work 126

Appendix A:

Appendix of Tables and Figures for Chapter 2 140

Appendix B:

Appendix of Tables and Figures for Chapter 3 178

Appendix C:

Appendix of Tables and Figures for Chapter 4 196

List of Tables

2.1 Sample Size for Community Survey . . . 16

2.2 Alternative Empirical Specifications - Choice of Instruments . . . 27

3.1 Sample Sizes for CEIP Participant Surveys . . . 51

3.2 Summary of First Stage Results . . . 75

4.1 Difference-in-Difference Coefficients . . . 105

4.2 Difference-in-Difference-in-Difference Coefficients . . . 115

A.1 Summary Statistics . . . 142

A.2 Summary Statistics for Social Network Measures . . . 143

A.3 Summary Statistics for Potential Instruments . . . 144

A.4 Variable Definitions . . . 145

A.5 Test Statistics Definitions . . . 146

A.6 Effect of Linking Contacts (*link*) on Employment (*hours worked*) . . . 147

A.7 Statistical Tests on the Effect of Linking Contacts (*link*) on Employment (*hours worked*) . . . 148

A.8 Effect of Bridging Contacts (*bridge*) on Employment (*hours worked*) . . . 149

A.9 Statistical Tests on the Effect of Bridging Contacts (*bridge*) on Employment (*hours worked*) . . . 150

A.10 Effect of Bonding Contacts (*bond*) on Employment (*hours worked*) 151
A.11 Statistical Tests on the Effect of Bonding Contacts (*bond*) on Employment (*hours worked*) 152
A.12 Effect of Network Size (*nwk*) on Employment (*hours worked*) 153
A.13 Statistical Tests on the Effect of Network Size (*nwk*) on Employment (*hours worked*) 154
A.14 Correlations between Social Network Measures 154
A.15 Effect of Linking Contacts on Hours Worked by Union Coverage 155
A.16 Effect of Bridging Contacts on Hours Worked by Union Coverage 156
A.17 Effect of Network Size on Hours Worked by Union Coverage 157
A.18 Effect of Linking Contacts on Hours Worked by Union Coverage and Age . 158
A.19 Effect of Linking Contacts on Hours Worked by Age 159
A.20 Effect of Linking Contacts on Hours Worked by Gender 160
A.21 Effect of Bridging Contacts on Hours Worked by Gender 161
A.22 Effect of Network Size on Hours Worked by Gender 162
A.23 Effect of Linking Contacts on Hours Worked by High School Completion . 163
A.24 Effect of Bridging Contacts on Hours Worked by High School Completion . 164
A.25 Effect of Network Size on Hours Worked by High School Completion . . . 165
A.26 Effect of Linking Contacts on Hours Worked by Skill Level 166
A.27 Effect of Link, Bridge, Bond, Tothelp, Nwk on Hours Worked under OLS . 167
A.28 Regression Statistics on the Effect of Link, Bridge, Bond, Tothelp, Nwk on Hours Worked under OLS 168
A.29 Effect of link, lnbridge, lnbond, lntothelp, lnnwk on lnhrswork 169

A.30 Regression Statistics on the Effect of link, lnbridge, lnbond, lntothelp, lnnwk on lnhrswork . . . 170
A.31 Heterogeneity of Impact of Social Networks (OLS) . . . 171
A.32 Heterogeneity of Impact of Social Networks (Log OLS) . . . 172

B.1 Summary Statistics . . . 182
B.2 Variable Definitions . . . 183
B.3 Frequency of Volunteering . . . 184
B.4 Volunteer and donor rates, population aged 15 and older, Canada, 2007 . . . 184
B.5 Volunteer rate by personal and economic characteristics, Nova Scotia 2007 . 185
B.6 Impact of Employment on Formal Volunteering: Probit in First Stage . . . 186
B.7 Impact of Employment on Formal Volunteering: LPM in First Stage . . . 187
B.8 Estimation of Probability of Employment: First-Stage Probit Results . . . 188
B.9 Estimation of Probability of Employment: First-Stage OLS Results . . . 189
B.10 Impact of Employment on Informal Volunteering: Probit in First Stage . . . 190
B.11 Reasons for Volunteering at Wave 2 (18-month) Survey . . . 191
B.12 Forms of Volunteering (in percentages) . . . 191

C.1 Classification of Decentralized and Non-Decentralized Sectors . . . 199
C.2 Summary Statistics of Aggregate District Expenditures . . . 200
C.3 Summary Statistics of Aggregate District Expenditures by Sectors - I . . . 201
C.4 Summary Statistics of Aggregate District Expenditures by Sectors - II . . . 202
C.5 Summary Statistics of Provincial Expenditures by Sectors . . . 203
C.6 Variable Definitions . . . 204
C.7 Difference-in-Difference Results . . . 205
C.8 Treatment Heterogeneity across Sectors - I . . . 206

C.9 Treatment Heterogeneity across Sectors II 207
C.10 Treatment Heterogeneity across Sectors (with Fixed Effects) I 208
C.11 Treatment Heterogeneity across Sectors (with Fixed Effects) II 209
C.12 D-D for Non-Decentralized Sectors (with Fixed Effects) 210
C.13 Difference-in-Difference-in-Difference (D-D-D) Results 211
C.14 Comparison of D-D for Decentralized and Non-Decentralized Sectors . . . 212
C.15 Probability of Reelection of District Nazim - Probit Estimation 213

List of Figures

2.1 CEIP Communities - Cape Breton Regional Municipality, Nova Scotia . . . 15

A.1 CEIP Timeline 141

A.2 Effect of Linking Contacts on Employment - Stata output page 1 174

A.3 Effect of Linking Contacts on Employment - Stata output page 2 175

A.4 Effect of Linking Contacts on Employment - Stata output page 3 176

A.5 Effect of Linking Contacts on Employment - Stata output page 4 177

B.1 Percentage of Program Group Members Actively Participating in CEIP, by Months from Enrolment 179

B.2 CEIP vs. Non-CEIP Employment - EI Sample 180

B.3 CEIP vs. Non-CEIP Employment - EI Sample 180

B.4 Timeline for CEIP Participant Surveys 181

B.5 Impact of Employment on Formal Volunteering: Stata output for Wave-2 . . 193

B.6 Impact of Employment on Formal Volunteering: Stata output for Wave-3 . . 194

B.7 Impact of Employment on Formal Volunteering: Stata output for Wave-4 . . 195

C.1 Pre-decentralization Local Government Structure 197

C.2 Post-decentralization Local Government Structure 198

Chapter 1

Introduction

This dissertation explores three different markets - the labour market, volunteering market, and the political market. In three separate but related studies, we investigate the effects of social networks on labour market outcomes in Chapter 2; we examine the determinants of volunteering behaviour in Chapter 3; and we estimate the impact of political, fiscal and administrative decentralization on public service delivery in Chapter 4.

None of these inquiries fits into the traditional mode of impact evaluation which measures the effects of policy and assesses changes in social welfare due to specific outcomes attributable to particular interventions. In all three cases, we utilize data from existing policy experiments in order to address broad questions of economic importance. The first two chapters use data from a labour market field experiment in Cape Breton Canada, while the third chapter exploits a natural experiment in budgetary allocations from a decentralization reform introduced in Pakistan.

This dissertation demonstrates uses of microeconometric estimation strategy with data generated by policy interventions and field experiments for examining key economic outcomes. These effects have otherwise proved hard to identify in the literature due to the problems to generate unbiased and consistent estimation of correlated unobservables and endogeneity. We exploit the fact that the policy experiments studied by us provide an exogenous source of variation in key outcomes of interest. We also utilize novel econometric approaches to ensure unbiased and/or consistent estimation of causal parameters. However, the contribution of this dissertation goes beyond mere estimation of policy effects as it focuses attention on the channels through which an intervention has an effect. Examining these channels matters both for enhancing our understanding of these markets, but also has significant policy implications (Deaton 1997). We thus explore not just what works, but why.

1.1 Key Research Questions and Potential Contribution

The importance of social networks in labour and non-labour market outcomes has been well documented in economics and sociology literatures for quite some time now. Less well known, however, is the impact of different types of social networks, especially of weak and strong ties. The first essay addresses two questions: Do social networks affect labour market outcomes? Is the effect, if any, differentiated by weak and strong ties? Overall the contribution of the first essay is to demonstrate using experimental data and current econometric techniques that weak ties matter in employment outcomes but strong ties do not. We also show that the effect of social networks varies along several interesting dimensions that include age, gender, union coverage, high school completion and skill level.

The second essay examines the effect of employment on levels of formal and informal volunteering. It fits into the empirical literature on the determinants of volunteering behaviour. It addresses the specific question: how does employment affect volunteering behaviour? Does it reduce the level of volunteering by squeezing the availability of free time, *or* does it motivate people through unobserved effects to volunteer more? We demonstrate that it does both, and that employment status has differential impact on levels of volunteering under different conditions.

Electoral decentralization and devolution of public service provision have become much advocated complementary pro-poor governance reforms in developing countries (Bardhan and Mookherjee 2006). But lack of exogenous variation makes it difficult to directly attribute the impact of decentralization on political accountability and on service delivery. The third study in this dissertation uses data from Pakistan to estimate the impact of decentralization. Despite having a healthy economic growth rate over the past sixty years, Pakistan's development experience has been marked by a conspicuously low level of human development for the majority of its citizens. Recent literature suggests Pakistan's poor social development outcomes as a consequence of governance structures that weakened political and bureaucratic accountability to citizens (Keefer et al., 2006). However, Pakistan underwent a major decentralization intervention from 2000 to 2001 that included political, fiscal and administrative reforms targeting these accountability failures. This makes Pakistan an interesting case for this study.

The contribution of the third essay is to demonstrate a large decentralization effect that is robust to the introduction of rigorous controls. It also shows an interesting pattern of treatment heterogeneity: the decentralization effect is only driven by a specific subset of sectors that does not include the social sectors. There is some evidence to suggest that this pattern is driven by the rational response of local politicians to electoral concerns and to the peculiar incentive structures introduced by higher levels of government.

1.1.1 Methodological Contribution

Traditional approaches to addressing the empirics of these questions have encountered several econometric problems as summarized in Manski (1993). These include the problems of correlated unobserved effects, endogeneity, misspecification and reflection. This dissertation demonstrates a strategy for estimating the impact of specific policy interventions covering the labour market, non-labour market and political market by using novel estimation techniques that control for these problems.

Chapter 2 examines the impact of social networks using two-step, efficient generalized method of moments (GMM) estimation in an overidentified context using cross-sectional data. Chapter 3 examines the effect of employment on volunteering behaviour using panel data in a non-linear model. It displays use of a control function approach to control for endogeneity. Chapter 4 uses difference-in-difference (D-D) and later difference-in-difference-in-difference (D-D-D) methodology to investigate the impact of decentralization reforms on the allocations for public services. We control for time trends, area fixed effects, and sector fixed effects. In our most extreme specification, we control for district-sector-time fixed effects but still get a large estimated decentralization effect.

In terms of estimation samples, Chapter 2 utilizes a cross-sectional sample from the community survey of the Community Employment Innovation Project (CEIP). Its identification relies on variation in social networks due to affiliation with the project in residents of the Program and Comparison communities. Chapter 3 utilizes a different participant survey from the same project that comprises four waves of interviews. Its identification relies on an experimental design with random selection of participants into Program and Control groups. Chapter 4 utilizes constructed data covering expenditure allocations from local governments in Pakistan prior to and after decentralization and exploits the fact that some types of expenditure were not decentralized and thus serve as controls in our estimation.

Chapter 2

Impact of Social Networks on Labour Market Outcomes

2.1 Literature, Research Questions and Methodology

That social networks matter in labour market outcomes has been known to economists and sociologists for some time, and in popular discourse for even longer, as expressed in the phrase: "it's not what you know but who you know!". However, our understanding of their impact on labour market outcomes remains limited. The theoretical predictions on how such networks affect outcomes is rather mixed on the direction of its predictions, whereas the empirical evidence faces numerous conceptual challenges. In particular, empirical researchers have to overcome problems associated with the issue that the formation of social networks reflects a variety of behavioural decisions. Since group formation must be treated as endogenous, this makes it particularly hard to identify the impacts of social networks as it is difficult to come up with a credible identification strategy with existing data sources.

This paper uses data from the Community Employment Innovation Project (CEIP), an innovative labour market field experiment recently conducted in Cape Breton, Canada. The experiment, because of its design, allows a source of exogenous variation in social networks for the participants that can be used to identify their impacts on employment.

2.1.1 Research Questions

Ever since Granovetter (1973) argued that weak ties are more important in finding jobs than strong ties, the strength of weak ties argument has been an open empirical question. For instance, Tassier (2006) examines available evidence on the question and concludes: "Given the intuitive appeal of the hypothesis, it is somewhat surprising that larger effects have not been found previously." This paper addresses this question by separately estimating the impact of strong and weak ties on employment outcomes. The primary research question is: do weak and strong ties have an effect on employment outcomes? How much does the effect of social networks vary if it is measured by weak versus strong ties? Does this effect vary along different dimensions that include gender, age, union coverage, high school completion, and skill level? And if it does, what accounts for this differential effect?

2.1.2 Literature

A network refers to a set of objects, or nodes, and a mapping of relations between them. More precisely, a social network refers to social relations among a set of personal contacts through which an individual receives support and information. This support depends on the number of contacts to which the individual is attached and on the strength of the relationship with those contacts.

Social networks play an important role in the labour market. The classic study by

Granovetter (1973) showed the importance of social ties, especially weak ties, in finding a job. Studies report that between 30 and 60 % of jobs are found through informal social network contacts (Ioannides and Loury, 2004). A number of studies provide empirical evidence of network-based job referrals and informational spillovers in the U.S. labour market (Bayer, Ross, and Topa, 2008; Topa, 2001). Networks are often thought to serve as a partial solution to information problems. For example, Montgomery (1991) argues that social networks can facilitate better screening of job applicants. Holzer (1988) suggests that it is the most efficient and the least costly job search method.

Calvo-Armengol and Jackson (2004, 2007) model transmission of job information through a network of social contacts and show *clustering* and *correlation* wherein employment and wages are positively correlated across networked agents both within and across periods. Another key feature of their models is *duration dependence*, wherein network status affects duration of employment

Fontaine (2008) explores a matching theoretical model where identical workers are embedded in ex ante identical social networks and studies the evolution of networks over time and characterizes the equilibrium distribution of unemployment rates across networks. He shows that networks induce new search externalities which shape the dynamics of the labor market.

Cahuc and Fontaine (2009) develop a simple matching model in which unemployed workers and employers can be matched together either through social networks or through more efficient, but also more costly, methods. In this framework, the use of social networks is an endogenous choice and decentralized decisions to utilize social networks in the job search process can be inefficient and give rise to multiple equilibria.

The size of the network is a salient factor determining how networks affect labour market outcomes. Munshi (2003) finds that Mexican migrants with exogenously larger networks have a higher probability of employment, as networks facilitate job search. He finds that the same individual is more likely to be employed and to hold a higher paying nonagricultural job when his network is exogenously larger, by including individual fixed effects in the employment and occupation regressions and by using rainfall in the origin-community as an instrument for the size of the network at the destination.

Beaman (2008) examines the random relocation of political refugees and sees significant differences in labor market outcomes based on the social setting that the refugees encounter. She reports that, depending on the vintage of other network members, having access to a larger network may actually lead to a deterioration of individuals' labour market outcomes due to competition among unemployed members for job information.

Laschever (2005) examines the random grouping of troops into military units in the United States World War I draft and finds that a ten percent increase in the average employment rate of a veteran's unit increases the veteran's employment rate by around three percent in expectation after correcting for other observables.

Another strand of literature explores how social networks affect crime, human capital accumulation and welfare participation. While it primarily focuses on non-labour market outcomes, this literature also underscores the indirect channels through which social networks may affect labour markets. For instance, Bertrand, Luttmer and Mullainathan (2000) empirically examine the role of social networks in welfare participation using data on language spoken at home to better infer networks within an area. Their results strongly confirm the importance of networks in welfare participation. Apinunmahakul and Devlin (2008) examine the link between social networks and private philanthropy and find strong evidence that networks promote donations of time and money.

2.1.3 Distinction between Weak and Strong Ties

There is no universally accepted definition of weak ties, but in general these refer to peripheral friends and random contacts who are not close in social space. Granovetter (1973) referred to weak ties as a network of acquaintances who are less likely to be socially involved with one another. We define a strong tie as one in which the relationship is repeated over time; for example, members of the same family or very close friends. We define a weak tie as one when social interaction between two persons is transitory. This, for example, includes random encounters.

In line with the distinction between strong and weak ties highlighted by Granovetter (1973, 1995), the literature differentiates between two types of networks (Putnam 2000). Granovetter emphasized that weak ties relay useful job information more frequently than strong ties. Others suggest that weak ties greatly increase the range of an individual's social network and thus allow them to have more information sources for jobs (Tassier 2006). Lin (1982) further emphasized the importance of weak ties and suggested that weak-tie job offers are drawn from a different (often superior) distribution.

Musick and Wilson (2008:p 285) emphasize the distinction between social ties that connect us to new people and social ties that reinforce our ties to people we already know - the distinction between "bridging" and "bonding" social ties, respectively. The distinction is made clearer by Smock (2004:66):

> "Bonding networks involve dense linkages among relatively small numbers of people. Each member of a bonding network typically knows every other member, and these relationships often overlap into multiple dimensions of the members' lives – like the members of a small town who attend church together, work in the same factory, know one another's families, and shop at the same

> stores ... bridging networks are composed of single-stranded ties that loosely connect large numbers of individuals. The relationships within these networks are generally less intimate or intense than those of the bonding networks ... members of bridging networks are typically linked to one another through indirect ties."

An important component of weak ties is its vertical dimension - vertical linkages in the network to people of higher status (or with broader networks) that can potentially give individuals capacity to leverage resources, ideas, and information to their benefit. Networks are also characterized by features like density (the extent to which individuals in a network know one another or share contacts), and heterogeneity (the extent of differences on various demographic and other characteristics e.g., religious, ethnic, charitable). It has been argued that less dense and more heterogeneous networks provide access to a greater variety of resources and allow for leveraging of ideas and opportunities into economic gain (Woolcock, 2001; Policy Research Initiative, 2003; Gyarmati and Kyte, 2003).

Strong ties may help find employment because of their more frequent contact with the individual, because these are more likely to know an individual is looking for a job, and because these are more motivated to providing job information than weak ties. Weak ties, although less frequent in contact and less specifically motivated to help, may have novel information sources and because of their increased range - there may be many more of these than strong ties (Tassier 2006).

2.1.4 Potential Contribution

The main challenge in estimating effects of social networks lies in identifying the causal effects of networks. This is because group formation is endogenous, and it is hard to tackle the “reflection problem” (Manski 1993) after group formation. These issues have been discussed in detail in a later section. Unbiased estimation of weak and strong ties has also proved hard in the literature because of lack of data and other challenges. The potential contribution of this study is that it identifies the impact of social networks by exploiting an exogenous source of variation in these. It provides consistent estimates of the effect of weak and strong ties in employment outcomes in labour market. It also displays interesting heterogeneity in social network effects by gender, union coverage, age, by skill and educational levels.

2.2 Description of Program and Data

The data for this study comes from the first wave of the Community Employment Innovation Project (CEIP) which was conceived by the Canadian federal government as a long-term research and demonstration project for testing an alternative form of income transfer payment in economically depressed areas through community involvement. The fundamental goal of the project was to improve the long-term well-being of workers in communities experiencing chronically high unemployment on the one hand, while contributing to the development of those communities themselves on the other hand. It was conducted by the Social Research and Demonstration Corporation (SRDC) and sponsored by Human Resources and Social Development Canada (HRSDC), while data collection was done by Statistics Canada.

The project involved two sets of linked interventions:

1. involving project sponsors and community groups in Program communities to plan and implement community projects by offering them free labour and other resources,

and

2. offering participants, randomly selected from welfare beneficiaries of the area, eligibility for 36 months of community work on projects developed by these sponsors/groups, with participants being paid salaries by the project in lieu of their welfare entitlements.

CEIP was set up in the Cape Breton Regional Municipality (CBRM) in Nova Scotia. Although individual participants were selected from across all of the CBRM, the community-based employment opportunities were concentrated in six local communities: Dominion, Glace Bay, New Waterford, North Sydney, Sydney Mines and Whitney Pier. Figure 2.1 shows the communities involved.

The selected (program) communities had to "volunteer" to participate in the project by means of a show of support by the majority of those attending public meetings held in each community. All six selected communities eventually chose to take part. They then had to go through a series of steps designed to engage members of the communities in the process of planning for and operating the projects that would employ the project participants[1]. Being "approved" meant that a sponsored community project was eligible to have CEIP participants assigned to work on it and to approve projects.

In total, these communities approved 295 projects submitted by different sponsors, which provided a total of approximately 1,300 positions and 2,100 unique placement opportunities. The number of work placements is different from the number of positions, as participants could work in multiple jobs over the course of their eligibility and several participants could fill the same job at different times. Over 250 community organizations were

[1]This involved the formation of a community board and its acceptance by the Project Implementation Committee, and preparation of a strategic plan; and following acceptance of its plan, a community board was authorized to begin approving projects submitted to it by organizations that wished to sponsor projects. Details of the process are given in Gyarmati et al. (2006, 2007, 2008).

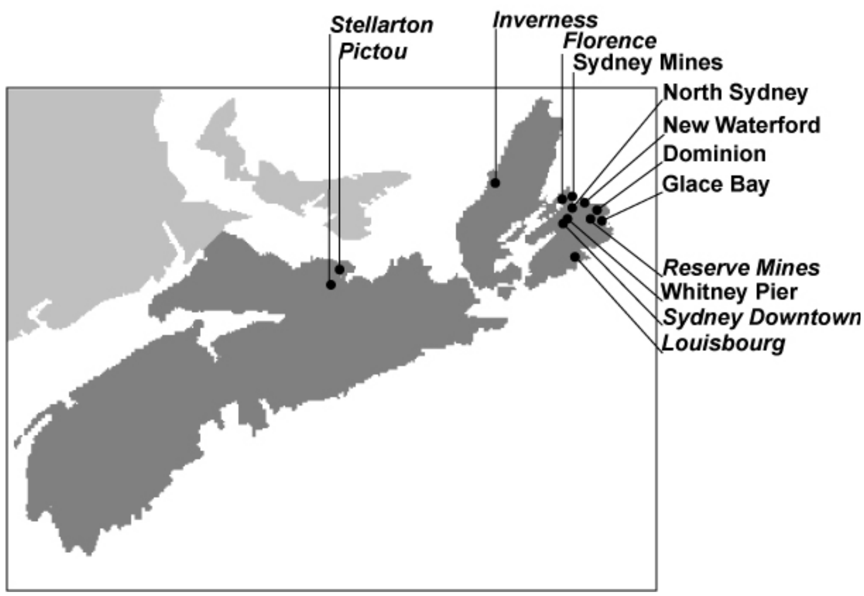
mobilized by program communities throughout the project period which began in the year 2000 and lasted up to the year 2005. These organizations developed projects that would employ participants who were paid by the project. The project made roughly 1200 full-time worker years and a few other resources available to these organizations/sponsors. The timeline of the project is shown in Figure A.1.

2.2.1 Data Sources, Sample Selection and Sample Sizes

The data set for this paper comes from the first wave of "Community Surveys Data" of CEIP. This was a three-wave longitudinal data set from a random sample of residents. However, this study uses wave-1 data only. The residents were drawn randomly by Random Digit Dialing (RDD) so that the data constitute a random sample of households in selected Cape Breton communities. The selected communities included six Program communities where the project-based activities of CEIP took place, and seven Comparison communities. The latter comprised a group of similar communities in Cape Breton and mainland Nova Scotia which were matched to the program communities that were to collectively serve as a counterfactual. The data was collected by the Institute for Social Research (ISR) at York University.

Comparison communities were selected on a high degree of similarity, measured by proximity score analysis, to program communities. The process involved the following steps: establishing a list of candidate communities; calculating pooled statistics for each of the descriptive community characteristics; calculating the squared Euclidean distance of the normalized Census characteristic variables from every other community; and finally selecting the comparison communities and community groupings with the shortest squared Euclidean distances (Gyarmati et al. 2008: appendix B).

The participants for the community survey were randomly selected from the Program and Comparison communities. The survey sample includes some individuals who were involved with the project (CEIP) in some capacity. These individuals are crucial for our identification strategy because their involvement with the project provided a source of exogenous variation in their social networks.



Note: The names of the comparison communities are shown in italics.

Figure 2.1: CEIP Communities - Cape Breton Regional Municipality, Nova Scotia

The data from the Wave-1 survey cover the following information:

- basic demographic data (e.g., age, gender, martial status, and education); health and activity limitation; satisfaction with life;
- employment and income data;
- data on time use and community participation; attachment and satisfaction with community, and migration; measures of cohesion, contact with neighbours, collective

engagement and trust;

- social network measures (described in detail below).

A list of relevant variables used in our analysis and their definitions is given in appendix Table A.4.

Table 2.1: Sample Size for Community Survey

	Program Communities	Comparison	Total
Sample Size	4,395	2,225	6,620

2.2.2 Summary Statistics

Summary statistics for the surveyed individuals are given in appendix Table A.1. The mean age of the participants in the estimation sample is 48 years; about 42 percent are males and 57 percent are married or living with a partner. In terms of family status, about 32 percent are single, without children; 11 percent are single with children; and 57 percent are couples with or without children. The mean household size is 2.68. In terms of the highest level of education achieved, 56 percent have completed high school; 9 percent have bachelor degrees; and 3 percent have some university education. About 23 percent have trade-vocational and apprenticeship diplomas.

The mean annual personal income is $24,000 while the mean annual household income is $39,000. About 40 percent of individuals have income levels below the Low-Income Cut-off (LICO) of Statistics Canada. About 25 percent of individuals receive pension from work, while 24 percent receive Employment Insurance (EI) or Social Assistance (SA). About 36 percent individuals are union members on their jobs, while about 11 percent individuals, though not covered by a union, have their wages covered by union contracts.

We also present summary statistics for both Program and Comparison communities. We find no significant differences over various outcomes of interest between these groups of communities.

2.2.3 Social Networks

The data set comprises different measures of social networks, entailing both weak and strong ties. We follow the definitions and proxies of measures of social networks that were adopted by the project (Gyarmati et al. 2008)[2]. One set of network measures covers home-based (or person-specific) networks: the number of family members and friends that the respondent sees and talks to; how many of these live in respondent's community, in Cape Breton or elsewhere in Nova Scotia; how many friends the respondent knew before 18 years of age; how many friends the respondent met for the first time in the last year; and how many family and friends know each other. Network size (*nwk*) is defined as the sum of the number of family and friends.

The other set of network measures is based on the kind of resources accessible from social contacts; these include bonding, bridging and linking contacts. The bonding network measure (defined as *bond* in the data set) is a sum of the following three variables:

1. Number of persons who can help the respondent with a 3-4 hour home project,
2. Number of persons who can help the respondent if sick with flu,
3. Number of persons the respondent can talk to if feeling down.

The bridging network measure (defined as *bridge*) is proxied by the following variable:

[2]Our results need to be qualified by the fact that these proxies may not be the perfect measures for social networks.

1. Number of persons who can loan the respondent \$500.

Total contacts that can provide various kinds of help (*tothelp*) is defined as the sum of bonding and bridging contacts.

Linking contacts represent contacts with persons of higher socio-economic status. Unlike other social network measures in our data that are measured through cardinal variables, linking contacts are proxied by a dummy variable *link* that takes the value 1 if both of the following dummy variables take the value 1:

1. Respondent personally knows a lawyer who can help; *and*

2. Lawyer is not a family member or relative.

Since weak ties comprise both linking and bridging contacts, in our empirical specifications, weak ties are proxied by *link* and *bridge* variables. Strong ties are proxied by *bond* variable.

Table A.2 display summary statistics for social network measures. These show that the mean number of family members and friends the respondent sees and talks to are 9.65 and 9.52 respectively. The mean number of bonding and bridging contacts are 24.2 and 4.53 respectively. About 30 percent of the respondent have linking contacts. There are no statistically significant differences between the Program and Comparison communities in these variables.

2.2.4 Estimation of Network Effects: The Challenges of Correlated Unobservables and Endogeneity

Estimation of social network effects presents a healthy set of challenges. Manski (1993, 2000) refers to a subset of these challenges as the "reflection" problem that arises when a researcher observing the distribution of behaviour in a population tries to infer whether the average behaviour in a group influences the behaviour of the individuals that comprise that group. Manski argues that network effects could be explained by three hypotheses: endogenous effects, exogenous effects and correlated effects. Endogenous effects arise when the tendency of an individual to behave in some way varies with the behaviour of the group; exogenous, or contextual, effects arise when the tendency of an individual to behave in some way varies with the exogenous characteristics of the group; and correlated effects arise when individuals in the same group tend to behave similarly because they have similar individual characteristics or face similar institutional environments. Manski (1993, 2000) explains that this distinction is important because endogenous effects generate a "social multiplier" through a feedback loop, whereas exogenous effects and correlated effects do not generate this "social multiplier".

Empirically distinguishing between exogenous, endogenous and correlated effects has proved to be challenging in the literature. Endogeneity can confound analysis of network effects because social networks can be influenced by the dependent variable or by some of the other factors that influence the dependent variable. Do people adjust their behaviour in response to that of people with whom they are connected, or do they choose their connections based on their behaviour (Jackson 2009)?

The problem of unobserved correlated characteristics arises because people tend to associate with others who have similar characteristics, and where some of those characteristics might not be observed by a researcher who can only condition on observed factors. This makes it hard to draw causal conclusions from the empirical data. For instance, if individuals commanding greater social networks are also the ones who share some common unobserved attribute, or shock, we might mistakenly attribute a correlation of outcomes among linked individuals to be due to the social link rather than to the unobserved attributes (Jackson 2009).

The empirical literature in economics has attempted to deal with these estimation challenges by taking advantage of timing, or by identifying instruments for peer behavior that are assumed to be plausibly exogenous. For example, Case and Katz (1991) and Gaviria and Raphael (2001) instrument for peer behavior using the average behavior of the peers' parents. Borjas (1992) regresses own behavior on measures of average human capital in the prior generation of one's ethnic group. Evans, Oates, and Schwab (1992) add an equation to explicitly model the fact that the teens in their data self-select into their peer group. Apinunmahakul and Devlin (2008) use information from the past behaviour of individuals (participation in team sports, youth group, etc. when young) as instruments for their social networks in later life. The studies by Munshi (2003), Beaman (2007), and Laschever (2007) that were mentioned above exploit having appropriate exogenously generated variation in the independent variables for identification.

Another approach is to use an experimental design where part of a unit is subject to a program that changes its behavior. This approach relies on having appropriate exogenously generated variation in the independent variables. Important papers that have used an experimental approach include Katz, Kling, and Liebman (2001); Ludwig, Duncan, and Hirschfield (2001); Sacerdote (2001); Duflo and Saez (2003); and Miguel and Kremer (2002).

2.2.5 Identification Strategy

The identification strategy of this paper relies on the fact that the CEIP project introduced a source of exogenous variation in the social networks of those who were affiliated with it. The project is likely to have changed bridging and linking social capital of some individuals by providing them greater opportunities of forming weak ties, or reducing the transaction costs to them of doing so. The mechanism in the project that potentially alters the social capital accessible by some community residents includes the process of community organization - meetings, canvassing, etc. It also includes the products of community projects - the delivery of new services like daycare or seniors' centres that bring diverse groups of people together.

Since there were few participating communities, those involved with the project often had to establish contacts outside of their hometown, giving them the opportunity to increase more distant contacts and enhance bridging social capital. Furthermore, those involved with the project may have developed links by meeting individuals, including project sponsors, training organizations and other participants, some of whom possessed extensive social networks and were in positions of influence[3].

This study demonstrates the importance of social network effects by exploiting an exogenous source of variation in social networks thereby eliminating the problem of network members selecting each other based on observable and unobservable characteristics. Exogenous variation implies that individuals' background variables are uncorrelated with social network measures. This allows us to measure the effect of size of different network measures on individual outcomes. Since the number of individuals whose social network measures were affected by the project is relatively small compared to the overall population, we do not expect the project to have large effects on project communities (Manski 1993, Heckman 2001).

In an ideal identification strategy, one would like to consider the complete network

[3]For paid Program group volunteers, the project also altered their social capital through the succession of assignments to community-based projects, but these volunteers, if randomly selected in the community survey, were excluded in order not to confound the effect of networks on employment.

structure of individuals, with not only data on the number of contacts of various types, but also a topography of who is linked to whom. Montgomery (1992) argues that the importance of weak ties can only be understood if one considers the entire network structure of individuals, subject some of them to an experimental variation, and examine the outcomes. Jackson (2009) suggests that the "reflection problem" can partly be overcome with more complete observation of the network patterns in a society, so that a given individual's peers can be directly observed and need not be inferred from the individual's own characteristics.

Because of the daunting nature of such a task, we limit our strategy to a reduced-form estimation of the impact of the size of different types of social ties on one feature of their employment outcomes, hours worked. This strategy is in line with most of the literature. In our opinion, this strategy is better suited for estimation than using proxies for network size, as used by Wahba and Zenou (2006) who use population density as a proxy for network size since their data set does not provide any information on the actual size of individuals' social network. A possible caveat to the external validity of results from this study, however, comes from the fact that the project was located in a region of chronic unemployment and relatively low incomes.

2.2.6 Estimation Strategy

The estimating equation for the reduced-form estimation used in this study is:

$$hrswork_{ij} = \alpha + \gamma N_{ij} + \mathbf{X}_{ij}\beta + \phi_j + \epsilon_{ij} \tag{2.1}$$

where i indexes individuals and j indexes communities. $hrswork$ is our outcome of interest (hours worked per week on all jobs). N_{ij} is a measure of network resources accessed by individual i in area j. X is a vector denoting control variables and ϕ_j denotes community fixed effects.

2.2.7 Instruments

Because of potential endogeneity of social network, we use instrumental variable estimation[4]. Some variables that can be used as potential instruments are:

1. Project-related:
   - Presently involved with CEIP
   - Heard about CEIP
   - Paid CEIP participant
   - Living in CEIP program community.
2. Not related to the project:
   - Respondent born in Cape Breton
   - Mother born in region
   - Father born in region
   - Number of years living in community
   - Number of years living at present address.

We capture involvement of individuals with the project in any capacity through the variable, *treat2*. This covers both paid and non-paid involvement with the project. Paid involvement can include the following:

[4]In fact, since the number of excluded instruments is greater than the number of endogenous regressors, we use two-step generalized method of moments (GMM) estimation.

- a Program group participant[5]who receives weekly wages (dropped from the sample),
- a worker paid directly by a sponsor of a community project,
- a paid employee of any of the organizations that took part in the project,
- or something else.

Non-paid involvement includes the following:

- control group members who were signed to become a study member, but not selected for the program projects, services or activities,
- volunteers to become a participant (awaiting outcome of random assignment to treatment),
- an unpaid volunteer for a sponsor of a community project,
- a member of CEIP Community Board or steering committee,
- attending one or more community meetings,
- an inactive participant (Program group member not in receipt of wages),
- or something else.

Since the outcome of interest for this paper is employment and the project Program group volunteers were offered employment eligibility, so program participation is not a

[5]Program group participants refer to the set of individuals who had been randomly selected from among welfare recipients and had later been randomly assigned into Program and Control groups. Only the Program group had been offered CEIP employment. This is briefly mentioned in Section 2.2 above, but is described in detail in Chapter 3 which relies exclusively on data from Program and Control groups for estimation.

valid instrument for these particular participants. This is because, although random assignment to program participation does produce an exogenous source of variation in their social networks, this also directly affects their employment due to the specific design of the project. As a consequence, for this particular group, the estimating equation merely picks the mechanical effect of program participation on employment. Therefore these participants (Program group members, 47 in number) are dropped from the estimation sample.

Summary statistics for potential instruments are displayed in Table A.3. These show that the mean proportion of survey respondents in our estimation sample who have heard about CEIP is about 23 percent; 2.4 percent are presently involved with CEIP while 0.6 percent are presently paid by CEIP. About 0.2 percent of respondents have other paid involvement with CEIP, whereas 0.2 percent have non-paid involvement with it. About 60 percent of the sample lives in CEIP Program communities.

Among non-project related potential instruments, the mean number of years lived by a respondent at the current address, in the community, and in Cape Breton is 18.4 years, 33.8 years, and 42.5 years respectively. About 86 percent of respondents were born in Cape Breton, whereas the fathers and mothers of about 75 percent respondents had been born in the region.

Finally, as one would expect, whereas we find no significant differences between the Program and Comparison communities on non-project related potential instruments, we find significant differences between these communities on project-related potential instruments. It is instructive to note that the distinction between these communities lies in the fact that, whereas Program communities received community projects, Comparison communities did not. However, those who were involved with CEIP in any capacity, whether as participants or in other paid or unpaid capacities, could belong to any community. Therefore, though we find a greater proportion of respondents from the Program communities connected in some way to the CEIP, we do find a significant number of respondents from the Comparison communities as well with some connection to the CEIP. In fact, there is a greater proportion of respondents from the Comparison communities who have non-paid involvement with the CEIP than there is from the Program communities.

It is important to highlight that we are relying on instruments to introduce variation in the social networks of our respondents that is exogenous to their current labour market behaviour. Our identification is coming from the group of respondents whose treatment status (social networks) change due to the instruments, in particular the ones reflecting their involvement with the project. This interpretation of the role of the project is in line with Johnson (2003) who finds that potentially beneficial but distant connections are often too costly to establish and maintain, and that the project (CEIP) can make such connections less costly. The project thus helped the participants to build bridging and linking social networks which, in turn, opened up access to labour market resources.

The CEIP community survey provides sampling weights for different communities and those have been used in the estimation.

Table 2.2 shows alternative combinations of instruments used in estimation. These instruments are defined as:

- **treat2**: Presently involved with CEIP
- **treat11**: Living in CEIP Program community
- **community2**: Number of years living in community
- **community4**: Born in Cape Breton region

- **community5**: Mother born in region.

Table 2.2: Alternative Empirical Specifications - Choice of Instruments

IV1	**IV2**	**IV3**
treat2	treat2	treat2
community5	community5	community5
community2	community2	
community4	community4	
treat11		

We include the variable 'living in CEIP program community' (*treat11*) as an instrument in specification **IV1**, but as a regressor in specifications **IV2** and **IV3**. Although we get similar results using all three empirical specifications, our preferred specification is **IV3** since it imposes the least restrictions on our estimation (see Table 2.2).

The number of observations in our overall estimation sample is 3,791 for the **IV3** specification, and 3,775 for **IV1** and **IV2** specifications[6].

2.2.8 Estimation and Identification Tests

Estimation is done by using the two-step efficient generalized method of moments (GMM) estimator which generates efficient estimates of the coefficients as well as consistent estimates of the standard errors[7]. The efficient GMM estimator minimizes the GMM criterion function $J = N * g' * W * g$. Here N is the sample size; g are the orthogonality or moment conditions which specify that all the exogenous variables, or instruments, in the equation are uncorrelated with the error term; and W is a weighting matrix. In two-step

[6]The difference in the estimation samples arises due to some missing values for $community2$ variable.

[7]We use the *ivreg2* command in Stata 10 with the *gmm2s robust* option. The various tests of identification described below are discussed in Baum et al. (2007), and in the Stata help file for *ivreg2* command.

efficient GMM estimation, the efficient or optimal weighting matrix is the inverse of an estimate of the covariance matrix of orthogonality conditions. The efficiency gains of this estimator relative to the traditional IV/2SLS estimator derive from the use of the optimal weighting matrix, the overidentifying restrictions of the model, and the relaxation of the i.i.d. assumption, as discussed in Hayashi (2000) and Baum et al. (2003, 2007).

We next conduct several tests of identification that are discussed below and are summarized in appendix Table A.5.

Tests of Overidentification

We conduct tests of overidentifying restrictions to test the statistical independence of the instruments from the error process. Since we use GMM estimation, we conduct a *Hansen J statistic* test, which is the GMM version of the *Sargan test*. This is known variously as the *Sargan statistic*, *Hansen J statistic*, *Sargan-Hansen J* test or simply a test of overidentifying restrictions. The literature highlights that this test should be performed as a standard diagnostic in any overidentified model[8].

This statistic is the value of the GMM objective function evaluated at the efficient GMM estimator $\hat{\beta}_{GMM}$. Under the null hypothesis that all instruments are uncorrelated with the disturbance process, it has a chi-squared distribution with degrees of freedom equal to the number of of overidentifying restrictions: the number of excess instruments (Baum 2006). The criticism of overidentification test by Altonji and Segal (1996) and Hall and Horowitz (1996) that it frequently over-rejects in small samples is not valid in our case since we have fairly large sample sizes.

[8]For instance, Davidson and MacKinnon (1993, 236): "Tests of overidentifying restrictions should be calculated routinely whenever one computes IV estimates."

We also test the orthogonality of a subset of instruments using the *C* or *difference-in-Sargan* test (Hayashi 2000, 218-221 and 232-234). This statistic is computed as the difference between two *J* statistics, the first of which is computed from the fully efficient regression using the entire set of overidentifying restrictions, while the second is that of the inefficient but consistent regression using a smaller set of restrictions in which a specified set of instruments are removed from the instrument list.

We also conduct test of exogeneity of the potentially endogenous regressors[9]. This test, robust to violation of conditional homoskedasticity, is defined as the difference of two Sargan-Hansen statistics: one for the equation with the smaller set of instruments, where the suspect regressor is treated as endogenous, and one for the equation with the larger set of instruments, where the suspect regressor is treated as exogenous. It is equivalent to a Hausman test statistic if homoskedasticity is assumed; see Hayashi (2000, 233-234). The null hypothesis is that the specified endogenous regressors can actually be treated as exogenous.

Tests of Weak Identification and Underidentification

We report tests of both underidentification and weak identification. The underidentification test is a test of whether the excluded instruments are relevant, that is, correlated with the endogenous regressors. Under the null hypothesis that the equation is underidentified, the test statistic is distributed as chi-squared. A rejection of the null indicates that the matrix is full column rank and the model is identified (Baum et al. 2007).

This test statistic is reported as the Kleibergen-Paap (2006) rank statistic. It can be seen as a generalization of underidentification tests to the case of non-i.i.d. errors, in which case the usual Anderson canonical correlation likelihood-ratio test is no longer valid. It is

[9]using the *endogtest* option in Stata 10.

reported as *idstat* in the estimated results, and its corresponding p-value is reported as *idp*.

Weak identification arises when the excluded instruments are only weakly correlated with the endogenous regressors. We report the robust Kleibergen-Paap Wald *rk* F statistic for weak instrument test (Stock and Yogo 2005), shown as *widstat* in the results.

We test the correlation of instruments with the included endogenous regressors by examining the fit of the first-stage regressions of the endogenous regressors on the full set of instruments, both excluded and included. This includes the squared partial correlation between the excluded instruments and the endogenous regressor in question, with the included instruments partialled out, as proposed by Bound, Jaeger and Baker (1995). Since we have a model with only one endogenous regressor, this statistic coincides with the related statistic proposed by Shea (1997), a partial R^2 measure that takes the intercorrelations among the instruments into account (Baum et al. 2007). The Shea's R^2 measure of instrument relevance is reported as *cd* in the results shown below. We also report the F-stat form of the Cragg-Donald statistic, shown as *cdf* in the results.

We report two statistics that provide weak-instrument robust inference for testing the significance of the endogenous regressors. The first is the Anderson-Rubin (1949) test statistic which is reported both as a Wald test and as a traditional F-test version (reported as *archi2* and *arf* below, with associated p-values reported as *archi2p* and *arfp* respectively). The second is the closely related Stock-Wright (2000) *S statistic*, which is a GMM-distance test (reported as *sstat*, with its associated p-value reported as *sstatp*).

The null hypothesis tested in both cases is that the coefficients of the potentially endogenous regressors in the estimating equation are jointly equal to zero and that the overidentifying restrictions are valid. The tests are equivalent to estimating the equation with the full set of instruments as regressors and testing that the coefficients of the excluded instruments are jointly equal to zero. Both statistics are distributed as chi-squared with degrees of freedom equal to the number of excluded instruments. Further details on these tests are provided in Dufour (2003), Chernozhukov and Hansen (2008), and Kleibergen (2007).

2.3 Results

We have described earlier the challenges of estimating the effects of social networks, and have discussed the reasons for why social networks cannot be considered exogenous in a simple model of labour market outcomes. We confirm this from our estimation sample, where a Hausman test rejects the exogeneity of social network measures. As a consequence, we resort to instrumental variable estimation. In this section, we first report the results from an unbiased instrumental variable estimation, and later report results from a biased OLS estimation.

An intuitive explanation of why OLS results are biased follows a similar logic to that employed by Card (1999) to explain why OLS generates biased results for the causal effect of education of earnings. Card's explanation is in terms of unobserved individual heterogeneity in the optimal schooling choice. In our case, possible sources of the unobserved individual heterogeneity are ability and motivation levels. At the background to our labour market effects lies individual choice of how many social networks to form, and this choice is governed by the benefits and costs of forming social links and the opportunities for forming these links. Estimation bias arises through the correlation between unobserved heterogeneity (ability/motivation) and the marginal cost of forming a link. If marginal costs are higher for individuals with more ability and/or motivation, and if these individuals would also tend to work more at any level of social network measures, then an application of the "omitted variables" formula to the employment equation entails a negative bias in estimation.

We do not report the complete first-stage results, but do show various statistics from the first stage. These include statistics from various tests of identification. The test statistics and their definitions are provided in appendix Table A.5. A specimen of full Stata output for both the first-stage and second-stage regressions for the model **IV3** of Table A.6 is shown in Figures A.2 to A.5.

Tables A.6, A.8, A.10 and A.12 show the estimation results for the regression of *link*, *bridge*, *bond* and *network size* on employment (*hours worked*), respectively. In each table, columns 1, 2 and 3 give estimation results using **IV1**, **IV2** and **IV3** as instruments for the endogenous social network variable, respectively. There are unreported additional coefficients, the same in each case, in all these tables. These include the respondents' occupational groups ($occ2$ up to $occ10$ with $occ1$ being the excluded category), the industry of main/last job ($ind2$ up to $ind10$ with $ind1$ being the excluded category), and the place of residence ($town2$ up to $town13$ with $town1$ being the excluded category).

Table A.6 displays instrumental variable regression output results for the effect of linking contact on employment. The effect of *link* is statistically significant at the 5% level for all three sets of instruments. The effect is also substantial in terms of magnitude and the positive sign corresponds with our prior expectation. Thus having a linking contact, a dummy variable, is associated with an increase in employment of about 18 hours per week. The coefficients and signs for the control variables are also in accordance with our expectation from previous studies. Thus age has a positive effect on employment, statistically significant at the 1% level, whereas its effect declines (age-squared is negative and also statistically significant at the 1% level). Educational status does not seem to have an effect on hours worked, when controlling for other factors. But the presence of younger children in the household has a significant negative effect on hours worked. Finally receiving a pension has a negative effect on hours worked, while being on welfare has a positive effect.

2.3.1 Robustness Checks

We conduct several robustness checks. Table A.7 displays a summary of statistical tests of identification concerning the effect of linking contacts ($link$) on employment (*hours worked*). The statistical tests based on the first stage regression results ($idstat$, $widstat$, and cd) suggest that the instruments are strongly correlated with the potentially endogenous regressor - *link* in this case.

The *Hansen J statistic* suggests that the instruments pass the exclusion restriction. For instance, the regression results attached as Figure A.2, pertaining to the empirical specification **IV3**, show that the *Hansen J statistic* has a value of 0.83 which implies a chi-squared with one degree of freedom p-value of 0.362 and thus constitutes no ground for rejecting the null hypothesis that all instruments are uncorrelated with the error term. We also test the orthogonality of a subset of instruments, *treat2* in this case, by using the *C statistic* and conclude that the suspect instrument can indeed be treated as exogenous.

We conducted similar robustness tests for other regressions, but their results have not been reported.

2.3.2 Effect of Bridging Contacts

Table A.8 provides estimates of the effect of bridging contacts ($bridge$) on hours worked (*hours worked*) under all three empirical specifications. Bridging contacts, measured as the number of persons who can lend the respondent $500, have a positive effect on working hours - one more bridging contact leads to an increase of about an hour of work in the preferred specification **IV3**. The effect is statistically significant at the 5% level for specification **IV3**, and significant at the 10% level for the first specification (**IV1**). The controls have similar effect as shown in the linking contact regressions. Table A.9 displays a summary of statistical tests of identification concerning the effect of bridging contacts on employment. The test results are similar to those in Table A.7 and provide evidence in support of our identification.

2.3.3 Effect of Bonding Contacts

We compare these results to the effect of bonding networks on hours worked, as shown by appendix Table A.10. Whereas the other variables retain their *ceteris paribus* effect, $bond$ does not have a statistically significant effect on hours worked in any specification. The results suggest that weak ties (as proxied by $link$ and $bridge$ in this data) have a significant effect on hours worked, whereas strong ties (as proxied by $bond$ in this data) do not have. Furthermore within weak ties, linking networks that signify connections to people of higher socio-economic status have a much greater impact on hours worked than mere bridging contacts, loose ties to other networks. Table A.11 displays a summary of statistical tests of identification concerning the effect of bonding contacts on employment. The test results are similar to those in Table A.7.

2.3.4 Effect of Network Size

When we measure network size as the aggregate of family and friends (nwk), its effect on hours worked turns out to be positive and statistically significant, as shown by

Table A.12. However, when we estimate the effect of network size as measured by the sum of $bond$ and $bridge$ variables (defined as $tothelp$ in this data), results (not shown here) show that its effect on working hours is not statistically significant. Table A.13 displays a summary of statistical tests of identification concerning the effect of network size on employment. The test results are similar to those in Table A.7.

A possible explanation of differential effects of the two measures of network size is that whereas our social network measures are proxies of actual social networks, these are also subject to differential measurement errors. Thus network size as the aggregate of family and friends (nwk) seems to suffer less from measurement errors than network size as the aggregate of $bond$ and $bridge$ variables ($tothelp$), each of whose component is a proxy for the kind of resources accessible from social networks and where there may be a considerable overlap between these proxies. Table A.14 displays the correlations between the individual components of nwk ($numfam$ and $numfri$) and $tothelp$ ($help1$, $help2$, $help3$, and $help4$[10]). The correlation between nwk and $tothelp$ is 0.64. The table suggests that there is a greater variation in correlations between the components of $tothelp$ than of nwk.

2.3.5 Network Effect by Union-coverage

Tables A.15 - A.17 display the estimation results of link, bridge and network size on hours worked by union coverage. Thus in Table A.15 the effect of $link$ on hours worked was separately estimated between those who have union coverage and those who do not. Union coverage in this work covers both individuals who are members of unions *and* those, while not union members themselves, whose wage contracts are determined by a union.

[10] $bond$ comprises the sum of $help1$, $help2$ and $help4$ while $bridge$ equals $help3$.

Similarly, Table A.16 displays the results of the effect of bridging contacts ($bridge$) on hours worked estimated separately for those with and without union coverage. Columns 1 and 2 display the results using instruments **IV1** with and without union coverage, whereas column 3 and 4 use empirical specification **IV2**, and columns 5 and 6 use empirical specification **IV3**. Table A.17 displays similar results for the effect of network size (nwk).

Table A.15 clearly shows that linking contact only has a positive effect on working hours for individuals without union coverage. This effect is of the order of 13-15 hours of additional work and is statistically significant at the 5% level in all empirical specifications. There is no effect of linking contact on individuals with a union coverage. A possible explanation is that the union coverage acts as the primary linking contact in the labour market for individuals with a union coverage, whereas individuals without union coverage require alternate linking contacts to access resources that can help them improve their labour market outcomes.

Table A.16 suggests that bridging contacts have a substantial effect, statistically significant at the 10% level, on individuals without union coverage. The effect is of the order of 0.7 to 0.9 hours. There is no statistically significant effect on individuals with union coverage.

Table A.17 suggests a similar story: network size is important in working hours for only individuals without a union coverage. This effect is small but positive and is statistically significant at the 5% level under all three empirical specifications. There is no statistically significant effect on individuals who are covered by a union.

Our results suggest a strong union effect on employment. We explore further whether the union effect may have been driven by an age effect because those without union coverage generally tend to be young while those with union coverage generally tend to be older individuals. However this hypothesis is not borne out by our data where there is a negative correlation of 0.18 between union coverage ($union$ dummy) and old age (dummy for age at or above 30 years). However, our results need to be qualified by the fact that only 8 percent of our estimation sample comprises individuals below thirty years of age.

We estimate the effect of $link$ on hours worked by union coverage separately for individuals below and above 30 years of age. The estimation results for the **IV1** specification are displayed in Table A.18. Each column in this table displays estimation results for a specific subset of individuals: the first column for individuals below 30 years of age who have union coverage, the second column for individuals below 30 years of age who do not have union coverage, the third column for individuals above 30 years of age who have union coverage, and the fourth column for individuals above 30 years of age who do not have union coverage. The effect of $link$ is statistically significant only in the second and fourth column - that is, the results are significant only for the set of individuals who do not have union coverage. Whereas the comparison may be hard because of there being few observations in the first specification, these results throw doubt on the effect of union coverage.

2.3.6 Network Effect by Age

We separately estimate the effect of social networks for different age groups. Table A.19 presents regression estimates for the impact of link on hours worked, separately estimated for individuals above and below thirty years of age. Columns 1 and 2 show results using **IV1** as instruments for individuals above and below thirty years, columns 3, 4 and 5, 6 similarly show combination of results using **IV2** and **IV3** as instruments, respectively. As mentioned earlier, these results need to be qualified by the fact that only 8 percent of our estimation sample comprises individuals below thirty years of age.

The results clearly show that linking social contacts have strong effects, statistically significant at the 5%, for individuals below thirty years of age, whereas linking contacts have no statistically significant effect on individuals above thirty years of age. Thus linking contacts are associated with an increase in hours worked of about 30 hours. Since this effect is much greater than the corresponding effect from the entire sample (about 18 hours), it is obvious that this age group is driving the results in the larger sample. A possible explanation is that younger individuals have lesser experience and thus perhaps stand more to gain from establishing linking contacts.

2.3.7 Network Effect by Gender

Tables A.20, A.21 and A.22 display social network effects disaggregated by gender. Table A.20 displays the results of linking contacts on hours worked by male and female gender status, Table A.21 displays male- and female-wise results of effect of bridging contacts on hours worked, while Table A.22 displays male- and female-wise results of the effect of network size on hours worked.

Table A.20 shows that linking contacts have a big effect on females, but not on males. Thus, whereas there is an effect of about 18 hours on females in two of the specifications, statistically significant at the 10% level in one of the specifications, there is no corresponding effect on the males.

Table A.21 shows a similar pattern - bridging contacts have a big effect on females, but not on males. Thus, whereas there is an effect of about 1 hour on females in all three specifications, statistically significant at the 10% level or better in all three specifications, there is no corresponding effect on the males.

Table A.22 again shows a similar pattern, where network size has an effect on hours worked for females (although the magnitudes are much smaller - 0.7 to 1.0 hour) in all three specifications, statistically significant at the 10% level, there is no corresponding effect on the males.

Our results suggest that linking and bridging contacts affect the labour supply for females more than for males. A possible explanation comes from the fact that whereas 88 percent of males in our sample have full-time jobs, only 70 percent of females are employed full-time. A greater number of female workers, compared to male workers, seem to be employed in part-time (casual) employment where hours of work are more flexible.

Tables A.20, A.21 and A.22 also suggest that the effect of having a greater number of children in the household varies starkly between males and females. Having more children has a generally positive effect on hours worked for males under most specifications. This effect on males is, however, not statistically significant under most specifications except in Table A.22 dealing with effect of network size. The effect of having more children on females is negative under all specifications. This negative effect is statistically significant in the regressions involving linking contacts at different levels of significance. This negative effect is also statistically significant at conventional levels under some specifications of the regressions involving bridging contacts and network size. These findings are in accordance with standard labour market results where the presence of greater numbers of children has generally a positive effect on the labour supply of males and a negative effect on the labour supply of females.

Another noteworthy finding is the contrasting effect of receiving pension payments on males and females. In the estimations on the effect of linking, bridging and total contacts, receiving pension has a strong negative effect, statistically significant effect at the 1% level, on the labour supply of males under all specifications. In contrast, receiving pension has no statistically significant effect on the labour supply of females.

2.3.8 Network Effect by High School Completion Status

Tables A.23, A.24 and A.25 display social network effects on hours worked disaggregated by high-school completion status for linking contacts, bridging contacts and network size, respectively, for all three empirical specifications. Table A.23 shows that linking contacts have a big effect in all empirical specifications of about 12-19 hours per week on individuals who complete high-school which is statistically significant at the 5% level in all three specifications. However, there is no corresponding effect on those who do not complete high school.

Table A.24 shows that bridging contacts have a small but statistically significant, at the 10% level, effect on individuals who complete high school in one specification, **IV3**. There is no statistically significant effect on high-school dropouts.

Similarly, Table A.25 shows that the network size has a small but statistically significant effect on individuals who complete high school in all three empirical specification. But there is no corresponding statistically significant effect on high-school dropouts.

2.3.9 Network Effect by High-Skilled / Low-Skilled Status

Table A.26 displays social network effects disaggregated by high-skilled and low-skilled status for the effects of linking contacts for all three empirical specifications. We define high-skilled individuals as those who either have a bachelors' degree ($college$), a university degree ($univ$) or have an apprenticeship or trade-vocational diploma ($diploma$). Low-skilled individuals are defined as those who have neither of the above-mentioned qualifications.

The results suggest that linking contacts have more effect on high-skilled workers as compared to low-skilled workers in terms of both magnitude and statistical significance (at the 10% level). There is no statistically significant difference on the effects of bridging contacts and network size between high- and low-skilled individuals; hence these regression results have not been included.

2.3.10 OLS Results

For comparison, we also run all the regressions in ordinary least squares (OLS). This is equivalent to treating network measures as exogenous. Table A.27 displays the OLS estimation results for regressions of $link$, $bridge$, $bond$, and network size ($tothelp$ and nwk) on hours worked. These variables retain statistical significance in OLS estimation under conventional assumptions. Thus *link* is statistically significant at the 1% level; $bridge$ is also significant at the 1% level; *bond* is not statistically significant; and $tothelp$ and nwk are statistically significant at the 5% level. These results correspond very well to the ones using GMM estimation. But while $tothelp$ was not statistically significant in GMM estimation, it is significant at the 5% level in OLS estimation under conventional assumptions.

Among the regressors, age measures ($ager$ and $agesq$), gender ($male$) and university education ($univ$) are all statistically significant at the 1% level. In terms of magnitudes, males and those with university education ($univ$) work more (about 5 hours each), age has a positive effect on working hours (about 1.2 hours) but at a declining rate ($agesq$ is negative, with a coefficient about 0.013). Receiving pension ($pension$) has a negative effect on working hours (from 2.1 to 2.8 hours in different specifications) and the effect is statistically significant at the 5% level in specifications involving link, bridge and network size. It is statistically significant at the 10% level in specifications involving bonding contacts and $tothelp$.

However, the major problem with OLS regressions is that these results produce biased estimation in severely underestimating the impact of social networks. For instance, whereas linking contacts ($link$) has a, *ceteris paribus*, effect of about 18 hours in GMM estimation, the corresponding effect in OLS estimation is only 2 hours.

A summary of the complete OLS regressions for $link$, $bridge$, $bond$, nwk, $tothelp$ differentiated by gender, union coverage, age, high school completion and skill level is presented in Table A.31. Each coefficient in the table comes from a separate regression, but where these other regression results have not been reported for brevity. Some results, for $bond$ and $tothelp$ specifically, have not been reported as these are not statistically significant.

2.3.11 Log OLS Results

We also run all the regressions in log-OLS form, where the dependent variable (hourswork) and the regressor of interest ($bridge$, $bond$, nwk, $tothelp$) have undergone the logarithmic transformation[11]. This log-log empirical specification is estimated by OLS and provides us an elasticity interpretation of the estimated coefficients on these variables.

[11] We do not apply the logarithmic transformation in our primary estimation reported earlier because, unlike for many outcomes of interest in the standard labour economics literature, we do not have a highly skewed distribution for our dependent variable, hours worked.

However, logarithmic transformation is not applied on one key regressor, $link$, which is a dummy variable. For this regression, we have a log-linear specification.

Table A.30 shows basic estimation results of OLS estimation in log specification form regressions. The dependent variable is in log form in all specifications, whereas the primary regressor of interest is in log-specification for $bond$, $bridge$, $tothelp$ and nwk, but not for $link$.

The results show that all measures of social networks - $link$, $lnbridge$, $lnbond$, $lnnwk$, $lntothelp$ - have significant economic effects on hours worked under logarithmic transformation ($lnhrswork$). The results are statistically significant at the 1% level for $link$, $lnbridge$, $lnnwk$, and $lntothelp$, and at the 5% level for $lnbond$ under conventional assumptions. The results also show that the elasticities are all highly inelastic (with values less than 0.1).

A summary of the complete log OLS regressions for $link$, $lnbridge$, $lnbond$, $lnnwk$, $lntothelp$ differentiated by gender, union coverage, age, high school completion and skill level is presented in Table A.32. Each coefficient in the table comes from a separate regression, but where the other regression results have not been reported for brevity.

2.4 Conclusion

This paper uses data from the Community Employment Innovation Project (CEIP) to estimate how changes in social network size and characteristics affect employment. The CEIP was a labour market experiment conducted in Cape Breton by HRSDC to test an alternative form of income transfer to recipients of welfare and EI. The results of this paper provide strong evidence that weak ties, as measured by linking and bridging contacts, matter in an important labour market outcome, hours worked, while strong ties, as measured by bonding contacts, do not matter. Thus both linking and bridging contacts have a substantial effect on hours worked which is generally statistically significant. However, bonding contacts do not have a statistically significant effect on hours worked. This suggests that the characteristics of a network have a key effect on the type of resources accessible from the network.

Our finding that the size of an individual's social network improves his or her employment outcomes is in line with the results in the recent literature in labour economics that utilizes instruments to estimate this relationship (Munshi 2003, Beaman 2008, Laschever 2005). However, the strength of weak ties hypothesis is still an open empirical question, as suggested by Tassier (2006). Our results provide some evidence of the importance of weak ties (bonding and linking contacts) and are in line with the results of Lin (1982) and Tassier (2006).

We later explore heterogeneity of network effects and examine the effects of social networks separately by gender, age, high school completion, skill level and union status. We find that the effects of social networks vary significantly between high-skilled and low-skilled workers, between high-school completers and dropouts, and by age, gender, and union status. Namely, the effects are statistically significant for individuals without union coverage, younger individuals (below 30 years of age), on females, on high-skilled individuals and on those who complete high school. We interpret these results as suggesting that social networks perhaps have greater effect on individuals who stand to gain the most from establishing these contacts.

The methodological contribution of this paper is that it deals with correlated unobservables and endogeneity of social networks in determining employment by relying on generalized method of moments (GMM) estimation and a randomized experiment. Comparison of estimation results from GMM and OLS reveal that OLS severely underestimate the impact of social networks. We exploit instruments from involvement with the project for providing an exogenous source of variation in social networks to come up with a consistent estimation that holds up in various tests of identification and appears to provide fairly robust results.

Other possible choices for exploring heterogeneity of network effects include industry structure, occupational and family characteristics of residents. Other relevant variables - like ethnicity, immigration status and country of origin - have not been included because of lack of data in the present data set. There is data on nature and degree of religious beliefs but there does not seem to be a major variation on these counts.

In future work, we wish to explore whether social networks have a significant impact on subsequent welfare participation of individuals. Social networks can affect individual behaviour towards welfare participation through two important channels: information and norms. Thus the interaction of disadvantaged with other disadvantaged persons may inhibit upward mobility as contacts may provide more information about welfare eligibility than job availability. This question will be examined by exploring the relationship between changes in social networks and changes in welfare participation. Changes in welfare participation will be explored by looking at the entry and exit of participants into welfare programs and by looking at their incomes from labour and welfare sources.

Chapter 3

Experimental Evidence on the Effect of Employment on Volunteering

3.1 Introduction

Volunteering as a form of altruistic behaviour is an important feature of social life. It is also an object of increasing interest in public policy in recent decades. It has been proposed as an alternative to "inefficient" government activity (Weisbrod, 1975). Isham et al. (1995) suggest volunteering as a means of ensuring the sustainability of public investments. Others have argued that since the opportunity cost of time is lower for the poor, volunteering can be used to self-target investments to the poor (Besley and Coate, 1992; Besley and Kanbur, 1993). Ostrom (1990) suggests that volunteering can be part of a solution to local collective action problems. Putnam (1993) and Fukuyama (1995) argue that volunteering nurtures civil society, builds trust, and can contribute to establishing accountable government and achieving high rates of economic growth.

As a consequence, understanding the determinants of volunteering behaviour has been an important question in economics and sociology literatures. Employment has been recognized as one of the key determinants of volunteering. However, despite its importance, it has been hard to estimate the effect of employment on volunteering because drawing such empirical evidence faces numerous conceptual challenges. In particular, empirical researchers have to overcome problems associated with the issue that the relationship of employment and volunteering must be treated as endogenous, since there are unobserved individual attributes that confound the causal link.

This paper uses data from the same Community Employment Innovation Project (CEIP). The project, because of its random assignment of individuals to employment eligibility, allows a source of exogenous variation in employment status of the participants that can be used to identify its impact on volunteering.

3.1.1 Research Question, Methodology and Potential Contribution

The primary research question of this paper concerns determining the effect of employment on volunteering levels. This question has important policy implications. Given the potential importance of volunteering, it is important to know the determinants of volunteering behaviour, which in turn can provide insights into the implications of labour market policy interventions on volunteering behaviour, both intended and unintended.

This paper defines volunteering as freely performing a job or providing a service outside the household without pay. It deals with formal volunteering which is done through an organization. Informal volunteering here is defined as helping other people directly, and not through an organization.

This paper makes a number of important contributions. It explores the determinants of volunteering behavior and in particular estimates the effect of one important determinant, employment. It provides consistent estimates of the impact of employment on volunteering behaviour. More importantly, it suggests that the effect is not simple and unidirectional, but a complex one mediated by many channels that link outcomes in the labour and the non-labour market. It thus highlights the need for examining more deeply the channels through which employment and volunteering appear to be linked.

Another contribution of this paper is methodological. It uses a control function approach for estimating the non-linear model, while bootstrapping the standard errors, and takes account of both unobserved heterogeneity and potential endogeneity by exploiting an exogenous source of variation in employment. A possible caveat on the external validity of the results from this study is that its estimation sample comprises predominantly low-income individuals from a specific geographic region in Canada.

3.2 Description of Program and Sample Selection

This chapter uses data from the Community Employment Innovation Project (CEIP). The project has already been described in Chapter 2, but this chapter uses a different data set from the project, one that focuses on the experimental part of the project. Participants for the experimental study were randomly selected from among welfare and EI beneficiaries residing in the Cape Breton Regional Municipality (CBRM). They were offered eligibility for 36 months of community work in return for foregoing their welfare and EI payments. The participants were paid at close to the minimum wages ($325 per week) for work on projects developed by the local community groups and approved by the local project selection board[1].

[1] The community wage was initially set at $280 per week, and increased over the course of the project to $325, in line with the increases in the provincial minimum wage (Gyarmati et al. 2008).

The selection was made from two broad groups - those from the Employment Insurance (EI) pool and those from the Income Assistance (IA) pool. Generally, those from the EI pool have greater links to the labour market, while those from the IA pool have weaker links to the labour market and greater levels of poverty. Participant selection and enrolment was carried out from June 2001 to June 2002.

Selection criteria for the participants reflected the rules and regulations that govern these transfer programs. Employment Insurance (EI) beneficiaries were randomly selected from a monthly derivative of the HRDC Benefits and Overpayments file which is used for administering EI claims and payments. Eligible Income Assistance (IA) recipients were selected from among IA recipients who expressed an interest in participating in the project after being notified by the Nova Scotia Department of Community Service (NS-DCS) about CEIP and their eligibility to participate in the project. Once selected, individuals were informed about the project. Those interested in participating in the study were required to complete an enrolment form consisting of an informed consent. They were then administered a survey that captured baseline measures on individual and socio-economic characteristics. The volunteers were then randomly assigned into Program (P) and Control (C) groups and both were surveyed at regular intervals.

The CEIP community projects covered the following sectors: health, safety and environment; services to youth; community upkeep and beautification services; arts and culture; services to seniors; recreational activities; services to seniors and others. The CEIP projects provided a range of occupations for participants throughout all 10 of the National Occupational Categorizations (NOC). Most of the jobs were service positions which include some skilled occupations, intermediate sales and service positions and a large proportion of elemental positions. The next largest categories of jobs were in business, finance, and administrative positions and natural and applied sciences (Gyarmati et al. 2007).

Those assigned to the Program group were offered eligibility for 36 months of employment on the project (CEIP) but were not mandated to work for the project throughout this duration. They were free not to take up the CEIP employment, to take absences from CEIP employment, and to work on non-CEIP employment and still retain their eligibility for CEIP employment till 36 months after random assignment. However, they were to lose their eligibility if they were to return to welfare as a major source of income. The typical participant worked on multiple projects in the social sector during the course of the project.

Figure B.1 shows the percentage of Program group members actively participating in CEIP (i.e., working for CEIP) by months from enrolment. It is obvious from the figure that the percentage of Program group members working for CEIP varied over the course of the project. Participation rates peaked for the EI sample at 78 percent during the fourth month after enrolment, and for the IA sample at 91 percent during the sixth month after enrolment, and declined gradually over the remaining follow-up.

Figures B.2 and B.3 display the proportion of CEIP versus non-CEIP employment for Program group participants by months from random assignment for EI and IA groups respectively. Figure B.2 shows that, among the EI sample, the proportion of Program group members employed full-time in non-CEIP jobs increased from about 10 percent in the first month after random assignment to 36 percent at 38 months. Figures B.3 shows that, among the IA sample, the rate of full-time employment in non-CEIP jobs was fewer than 10 percent for most of the eligibility period.

The time-line of the project is shown in appendix Figure B.4. It shows the timing for random assignment, job placement, baseline survey and for follow-up surveys.

3.2.1 Data Sources and Summary Statistics

Our primary data source for this study comes from CEIP participant surveys, conducted at baseline with followup surveys at 18 months, 40 months and 54 months after random assignment[2]. Other sources of data include the Program Management Information System (PMIS) database and some administrative data. PMIS data was used in conjunction with the follow-up survey data to derive employment and earnings data for project participants. Table 3.1 displays sample sizes for the various surveys.

Table 3.1: Sample Sizes for CEIP Participant Surveys

Wave	Program	Control	Total
Baseline	757	757	1514
18-month	707	656	1363
40-month	651	611	1262
54-month	599	553	1152

The data, which had to be extensively cleaned before estimation by us, include the following information:

- basic demographic data (e.g., age, gender, and martial status, education and training); health and activity limitation; satisfaction with life, etc.,
- employment data covering job characteristics such as industry and occupation classification, job duration, absences, pay rate, seasonal or non-seasonal, characteristics of employer, and unionization,
- data on personal and household income and benefit payments,

[2]Each survey was staggered over time since induction into the project lasted over many months. The 54 month survey was conducted at least 12 months after the program ended for all participants.

- data on volunteering (described in more detail below).

Summary statistics showing characteristics of participants at baseline are displayed in appendix Table B.1. Participants typically lived in households composed of two or more persons at baseline. They were likely to have lived in Cape Breton for all their life. The vast majority reported being in good health at baseline. Linguistic and ethnic variables were not a part of the surveys and the area demographics show that there is little variation on these counts in any way.

There are no significant differences between the Program and Control groups in the listed variables from t-tests. Participant samples mostly represented disadvantaged populations, but there is considerable variation along several dimensions. Thus the Employment Insurance (EI) sample is more likely to be male, at 58 per cent, while 62 per cent of the Income Assistance (IA) sample is female. The EI sample is typically older, with an average age of 40, while the IA sample age was 35 at baseline. The EI sample had a higher educational attainment, with 69 per cent holding a high school diploma compared to 60 per cent of the IA sample. The household income for most EI sample members was under $30,000 during the 12 months before enrolment, while the household income of most IA enrollers was less than $20,000 with over half of the sample reporting income of less than $10,000. The EI sample had a longer work history than IA sample members at baseline. They were, however, also more likely to be unemployed due to a layoff, contract termination, or because their employer moved or closed down.

3.2.2 Volunteering

Since there are no data on charitable donations in our data set, our primary measure of volunteering is frequency of formal volunteering. Formal volunteering activities in our surveys can take the form of working as an unpaid member of a board or committee, canvassing, supervising activities, teaching, or delivering food on behalf of an organization. The participants were asked how often they participated in unpaid volunteer activities for groups or organizations in the last one year. The measure, represented by $forvol$ in our data set, records the responses on the following scale:

1. never
2. less than once a month
3. once a month
4. once a week
5. few times a week
6. everyday.

Table B.3 displays variation in levels of volunteering between Program and Control groups over time. It is worthwhile to explore how provincially and nationally representative our sample is. Tables B.4 displays volunteering and donor rates for the population aged 15 and older from the 2007 Canadian Survey on Giving, Volunteering and Participating (Hall et al. 2009). Table B.5 displays the volunteering rate and distribution of volunteer hours from the same source, by personal and economic characteristics, for the population aged 15 and older from Nova Scotia. Nova Scotia performs better than the national average in terms of volunteering: the volunteering rate in Nova Scotia is 55% compared to a national average of 45%, the average hours volunteered in Nova Scotia are 183 per year compared to a national average of 166 hours. However, whereas the volunteering rate in Nova Scotia increased from 48% in 2004 to 55% in 2007, the average annual volunteer hours dropped from 195 hours in 2004 to 183 hours in 2007 (Hall et al. 2009).

A comparison of these tables and the summary statistics mentioned above reveals that our sample is more representative on some dimensions than others. Thus the average volunteering rates in our sample fall between the national and the provincial rates at baseline. Our sample appears to be provincially representative on many dimensions. However, since our sample comprises low-income households, there is far less variation in household income than is there at the provincial level.

3.3 Literature on Volunteering

The theoretical economics literature on volunteering is extensive. Important papers include Arrow (1974), Becker (1974), Rose-Ackerman (1982), Sugden (1984), and Andreoni (1998) among others. There are also several empirical studies which seek to explain the determinants of volunteering. These include Menchik and Weisbrod (1987), Freeman (1997), Vaillancourt (1994), van Dijk and Boin (1993), and Day and Devlin (1996). The recent literature on the subject includes important papers by Apinunmahakul and Devlin (2008), Prouteau and Wolff (2006), Apinunmahakul et al. (2009). The literature identifies a number of factors that influence volunteering behaviour. These include gender, education, income, employment and community characteristics.

Economists have mainly focused on investment or consumption motives to explain why people undertake volunteer activities but there is a recent literature that emphasizes the relational motive to undertake volunteer activities, a concept previously emphasized by social psychologists. Volunteering in this view is seen as a way to build friendly relationships.

Employment plays an important role in volunteering, but the direction of its effect on volunteering behaviour is not ex ante obvious. According to one view, employment has a negative effect on volunteering because it reduces the free time available to individuals that could be used for volunteering. In other words, this view states that work time squeezes out volunteer time.

Furthermore, the time allocation/household production model assumes that an individual's (household) time can be used not only in leisure activities or in the workplace but also in productive, non-market activities including volunteering. Since volunteer work imposes costs in the form of income foregone because of absence from paid work, it is natural to assume that if actors are rational then the more they earn through working, the less they will volunteer. Wage rate effects operate through the opportunity cost of time.

However, employment does not necessarily have a negative effect on volunteering and in fact the two may be positively related through individual attributes that are mostly unobservable to outside researchers. Employment status imposes extra time demands on people, but also provides opportunities to them to socialize in their communities. Musick and Wilson (2008) suggest that people's time can be elastic if they are sufficiently motivated to take on a number of tasks, and that having a paid job increases the likelihood of individuals learning about volunteer opportunities or being asked to do volunteer work. Volunteering also provides access to social networks which, in turn, can enhance employment opportunities, as demonstrated by our previous chapter.

Another way of thinking about working and volunteering is to treat volunteer work as an unpaid productive activity that is costly to perform. Even if volunteer work itself is not expensive, it is often channeled through voluntary organizations that expect their members to pay dues and incur other incidental expenses. "Ability signaling" may provide another link between work and volunteering: status signals associated with higher incomes earned through more work mean wealthier people are more likely to be the target of recruiters for volunteer work.

Volunteer work can also provide an alternative training ground for learning job skills. Day and Devlin (1998) suggest that volunteering may aid in the acquisition of marketable skills and business contacts or may serve as a favourable signal to employers. Formal volunteering activities are more likely to build human capital than informal volunteering because they occur in an organizational setting and are therefore more likely to provide opportunities to develop work-related contacts and work experience (Gyarmati et al. 2007). For instance, 47 per cent of Canadians cite networking or meeting people as a motivation for formal volunteering, and 22 per cent want to improve their job opportunities (Statistics Canada, 2006).

Finally, working may also affect preferences of individuals. Employed people may have a greater “stake” in a number of issues that call for volunteer work. Thus they may come to have a greater taste for social issues like safer neighborhoods and better schools. These arguments imply a positive association between work and volunteering: the more we work and earn, the more we tend to volunteer.

3.3.1 Evidence on the Supply of Volunteer Labour

There is some survey evidence in support of the argument that more work time means less time available for volunteering, and that higher wages are associated with more labour supply and less volunteering due to higher opportunity cost of time. People commonly cite shortage of time as a reason for not volunteering, seeing volunteer work as too demanding a commitment of time from them (Musick and Wilson 2008). In one study, 30 percent of the respondents blamed long work hours, odd work shifts, and frequent work-related travel for their failure to volunteer (Profile of Illinois 2001). Smith (1998) reports from a United Kingdom survey that 58% of respondents who were not volunteering said they did not have the time. Hall et al. (1998) report that three-quarters of Canadians surveyed gave lack of time as the main reason for not volunteering more.

While self-reported data may not be entirely credible, there is also evidence to suggest that time constraints are actually mentioned by people who are most likely to be short of time. The Canadians most likely to say they cannot volunteer because they lack the time are generally young to middle aged adults working full-time (Hall et al. 2001). Kohut (1997) reports that almost twice as many employed Americans wish they had more time to volunteer than unemployed and retired persons. Data also suggests that volunteers who try to combine volunteer work with paid work feel more pressured. Goss (1999) found that volunteers feel more "hassled" than non-volunteers. A Canadian survey found that volunteers were more likely than non-volunteers to say their work hours are too demanding (Gomez and Gunderson 2003).

On the other hand, there is also evidence to suggest that people with jobs may volunteer more than people without them. This suggests that volunteering behaviour may be related to employment through unobserved attributes like motivation and ability. For instance, volunteers may potentially be highly motivated and energetic people who can find a way to fit both working for pay and volunteering into their daily schedule (Musick and Wilson 2008). This is supported by evidence that teenagers who have part-time jobs while in school are more likely to volunteer than those without a job (Sundeen and Raskoff 1994). It may also be the case that people with children to support are more likely to work, and that the children may get them to volunteer more - for instance, in school sports activities, or church involvements.

There is strong evidence to suggest that part-time employees volunteer at a higher rate than people without jobs (Johnson et al. 2004; Robinson and Godbey 1997; Vaillancourt 1994). This is also suggested by the 2000 Canadian Survey on Volunteering (Lasby 2004). The 2005 Current Population Survey (CPS) of volunteering in the United States found that 38.2 percent of part-time workers had volunteered in the past twelve months, compared to 29.8 percent of full-time workers, and 24.4 percent of those not in the labour force (U.S. Bureau of Labor Statistics 2005).

However, many correlational studies indicate that full-time work is an impediment to volunteering. These studies control for other factors that also influence volunteering - such as education, marital and parental status, and physical health. Ravanera et al. (2002) report for a study of Canadian data that controlling for these variables, full-time workers volunteered at a *lower* rate than those not in the labor force. Oesterle et al. (2004) report similar results: each additional month spent working full-time lowers the odds of volunteering by 4 percent. Wilson and Musick (1997) report, that among those who worked forty hours a week or more, the relation between work hours and volunteer hours was positive. The tendency for people who spend fifty or sixty hours a week at work to volunteer more than people who work the standard forty-hour week contradicts the time squeeze theory. Freeman (1997) attributes this to differences in taste, ability, and energy between people who work the standard work week and people who regularly work longer hours.

Day and Devlin (1998) use a Canadian data set and test the hypothesis that volunteer work acts as an investment in human capital and increases an individual's earnings. Their results suggest that volunteer work does indeed increase individuals' earnings and estimate the return to volunteering to 6-7 percent of annual earnings. Prouteau and Wolff (2006) focus on the investment motive for volunteer work and examine whether volunteer work has economic payoff upon the labour market in France. Their findings are more consistent with a consumption motive and they suggest that volunteering is carried out with a relational purpose.

Taniguchi (2006) examines gender differences in the effects of employment on volunteering and reports that there is a statistically significant difference in the way employment status affects men's and women's volunteering behaviour. He finds that relative to full-time employment, part-time employment encourages women's volunteer work but not men's, while unemployment exclusively inhibits men's volunteering. Hackl et al. (2007) explore the consumption and investment motives for volunteering and find strong statistical evidence for the investment model with a significant impact of volunteering on the wage rate. They explain that the number of volunteering hours plays a major role in explaining this wage premium and attribute this finding as supporting the significance of skill acquisition to accumulate human capital, the importance of deepening of social contacts and signalling willingness to perform.

Prouteau and Wolff (2008) examine data from a French survey on association membership and volunteer work and report evidence supporting the relational motive. Their results suggest that working as a volunteer in an association has a causal impact on the probability of making friends in that association. Antoni (2009) investigates the intrinsic and extrinsic motivations to volunteer and by analyzing how participation affects social network formation. He suggests that intrinsic motivations enable people to create relations characterized by a significant degree of familiarity while extrinsic motivations promote the creation of networks from a quantitative point of view, but they do not facilitate the creation of relations based on a particular degree of confidence.

Apinunmahakul et al. (2009) explore the relationship between individual contributions of time and money and the paid labour market. In particular, they examine the question whether people who work in the paid labour market behave differently than those who do not when it comes to their private philanthropy. They report that, for employed individuals, the donations of time and money are largely complementary to each other, and that employment status as well as gender affect how individuals respond to fiscal policies.

In summary, the determinants of volunteer labour supply is an active area of research and is an open empirical question. In theory, the effect of work hours on volunteering can go either way. It is also hard to estimate an unbiased or consistent effect because of unmeasured attributes of individuals involved. In the next section, we describe the economic model underlying our analysis, and later we describe how we overcome these estimation hurdles.

3.4 Model

Following Menchik and Weisbrod (1987), we model the volunteering activity of an individual who faces an exogenously determined wage rate, w. The individual is free to adjust his or her work-time, leisure, and volunteer time in accordance with exogenously-determined non-earned income, preferences, and the prices faced. Our analysis of volunteering behaviour concentrates on the supply side of volunteer labour. The variables used are:

Predetermined variables:

w = wage rate for market work

T = endowment of available time

y = nonearned (nonlabor) income

Endogenous variables:

t_l = time spent on leisure - that is, time not used for market work or volunteer activities,

t_v = time spent on volunteer activities,

t_m = time spent on market labor ($t_m + t_v + t_l = T$),

C = conventional consumption expenditures.

In this model, the income concept is 'full' or 'potential' income - that is, $wT + y$.

The individuals are assumed to attempt to maximize their utility functions, assumed to be quasi-concave and increasing in all goods, subject to a budget constraint. The fact that voluntary work is time consuming implies that there is an opportunity cost involved with volunteering. The income constraint is defined by the product of the wage rate w and the working hours, and non-earned income. Hence, by providing an additional hour of volunteering individual full income is affected. If the wage rate changes the allocation of time and, therefore, income will change as well. An increasing (decreasing) wage rate will be associated with a decline (increase) of voluntary work due to the substitutional relationship between paid work and volunteering.

The Lagrangian function, L, is:

$$L = U(t_l, t_v, C) + \lambda[(w(T - t_l - t_v) + y) - C] \tag{3.1}$$

The first-order conditions yield:

$$
\begin{aligned}
&\text{(i)} \quad \frac{\partial U}{\partial t_l} - \lambda w = 0, \\
&\text{(ii)} \quad \frac{\partial U}{\partial t_v} - \lambda w = 0, \\
&\text{(iii)} \quad \frac{\partial U}{\partial C} - \lambda = 0, \\
&\text{(iv)} \quad [(w(T - t_l - t_v) + y) - C] = 0.
\end{aligned} \tag{3.2}
$$

from (i) and (ii): $\frac{\partial U}{\partial t_l} = \frac{\partial U}{\partial t_v}$.
from (iii): $\lambda = \frac{\partial U}{\partial C}$.
from (iv): $t_v = T - t_l - \frac{C-y}{w}$.

This suggests that the marginal utility of consumption, the increase of level of utility owing to the last dollar spent on consumption, must be the same that is obtained of the last dollar that is not earned due to the time spent on leisure or voluntary work. The implication of this model is that the individuals will allocate time to volunteering until its returns equal, at the margin, the returns from time spent on leisure, while the opportunity cost of both is the wage rate.

We linearize the last equation to get an estimating equation.

$$t_v = a + bt_m + \mathbf{c}\,\mathbf{Y} + \nu \tag{3.3}$$

where t_m is the time spent on non-leisure activities ($T - t_l$), and $\mathbf{Y}$ is a vector containing wages, income, consumption etc.

Since we do not have data on the precise time spent on labour, we proxy it with employment status, job, a dummy variable which is equal to one if the respondent reported being employed full-time at the time of the survey (corresponding to 35 or more hours of work per week). In terms of our theoretical set up, this variable takes the following values for each individual for each survey:

$$job = \begin{cases} 1 \text{ if } t_m \geq 35 \\ 0 \text{ otherwise} \end{cases}$$

As discussed earlier, volunteering in our data set is measured by a self-reported variable, $forvol$, that measures how often the respondent volunteered during the last 12 months. This is reported on an ordered response scale that takes six values - 0 (never), 1 (less than once a month), 2 (once a month), 3 (once a week), 4 (few times a week), and 5 (everyday). This suggests the use of an ordered response model. Accordingly, we use an ordered probit model for estimating the effect of employment on volunteering levels.

This leaves us with a non-linear ordered response model with a potentially endogenous explanatory variable, and calls for novel approaches to control for this potential endogeneity[3].

3.5 Empirical Strategy

Estimation of the impact of employment on volunteering is confounded by the potential endogeneity of employment which may occur for a number of reasons. For instance, volunteers may be high-ability persons who do better in school, get better educational credentials, and, as a result, get better jobs. It may also be that volunteer work is a form of

[3]We first attempted a linear model of 2SLS estimation using the Local Average Treatment Effect (LATE) approach of Angrist and Imbens (1994, 1995). However, transformation of the ordered response variable $forvol$ into a binary variable resulted in a substantial loss of information and therefore we abandoned the approach in favour of a non-linear model using a control function approach but with a similar identification as the one used in the LATE approach.

occupational training in which important job skills are learned, which in turn makes it easier to compete for good jobs. Another source of endogeneity is that volunteer work forges social ties with a wider network of socially heterogeneous people, which, in turn, enhances job prospects by increasing the chances of learning about, or having personal contact with someone who is offering, a desirable job. We provide some suggestive evidence of this in Chapter 2 of this dissertation.

Our identification strategy relies on the fact that the project (CEIP) offered an exogenous source of variation in employment for some participants. This is because the project randomly assigned participants to Control and Program groups. The project first randomly selected beneficiaries of Employment Insurance and Income Assistance programs, offered them participation in the project, and from those who accepted the offer, it randomly assigned half of them to the Program group which was given employment eligibility for 36 months, while the Control group was not offered anything. Both groups were, however, surveyed at regular intervals. Randomization was conducted by Statistics Canada on dedicated random assignment software application, and procedures were adopted to protect the integrity of the process[4].

The potential endogeneity of employment calls for the use of appropriate instrumental variables for consistent estimation controlling for the endogeneity of employment. Random assignment to treatment can thus serve as a good instrument since it is not correlated with the disturbance process in our volunteering equation, because it was assigned randomly. But it is likely to be highly correlated with the employment status of survey respondents. Random assignment cannot be perfectly correlated to the employment status of participants because the encouragement design nature of the project intervention results in partial

[4]The procedure adopted and the different checks applied are described in detail in Greenwood et al. (2003, p.121-123).

compliance, as suggested by Figures B.1, B.2 and B.3 that have been discussed earlier.

Our outcome of interest, $forvol$, is an ordered response variable with the values assigned to each outcome conveying ranking and have ordinal meaning. However we cannot infer from these values that the difference between five and two is somehow thrice as important as the difference between one and zero. If our model contained only observable exogenous regressors, we could use a standard ordered response model for estimation. However, we have argued earlier that employment and volunteering are jointly determined and thus employment cannot be considered as exogenous in our model. In the results section, we will conduct a Hausman test that will present convincing evidence against the exogeneity of employment. This endogeneity of employment adds an additional layer of complexity to our estimation over the standard ordered response model. We deal with this endogeneity by using a control function approach. In a simple version of the control function approach (Imbens and Wooldridge 2009, Terza et al. 2008), we estimate auxiliary regressions in the first stage to generate first-stage residuals, and then use these residuals as additional regressors in a second-stage estimation.

Imbens and Wooldridge (2009) describe that nonlinear models with endogenous explanatory variables are typically estimated by either plugging in fitted values from a first step regression (in an attempt to mimic two-stage least squares (2SLS) in linear model); by maximum likelihood; or by a control function approach, the approach that we adopt. They suggest that certain nonlinear models with endogenous explanatory variables are most easily estimated using the control function method, which relies on the same kinds of identification conditions as the standard instrumental variables methods - the 2SLS or GMM.

The control function method involves determining the projection of the endogenous explanatory variable onto the exogenous variables and adding the error term in the reduced form equation for the endogenous regressor in the structural equation for the outcome of interest in order to control for endogeneity. Since we do not observe this error term from the projection of the endogenous regressor onto the exogenous regressors, we can consistently estimate it in a first step regression and include the residuals in the structural equation and find consistent results. In a linear model, 2SLS also provides a method of consistent estimation (Greene, 2003). Imbens and Wooldridge (2009) maintain that for a linear model, the estimates from a control function approach are *identical* to the 2SLS estimates.

But this consistency property of 2SLS in the fully linear models does not generally extend to non-linear models. Furthermore, the identity between the estimates from a control function approach and a 2SLS approach[5] also breaks down in non-linear models. This is especially true for multinomial and ordered response models involving categorical endogenous variables[6]. Use of control function approach allows one way of estimating such models.

Terza et al. (2008) show that, for a non-linear model, a specific control function method, which they name as **two-stage residual inclusion** (2SRI), generates consistent results in a generic parametric framework, whereas 2SPS does not. The 2SRI estimator can be cast as a special case of the conventional generic two-stage optimization estimator, as discussed in Newey and McFadden (1994), White (1994, Chapter 6), and Wooldridge (2002, Chapter 12).

Ignoring the endogeneity issue initially, we represent all regressors, both exogenous as well as endogenous, with the vector **X**, and derive an ordered probit model for $forvol$ (conditional on explanatory variables) from a latent variable model. We assume that a latent variable $forvol^*$ is determined by

[5] The non-linear version of 2SLS is also sometimes named as 2SPS - two-stage predictor substitution.

[6] "Allowing endogenous explanatory variables (EEV) in multinomial response models is notoriously difficult, even for continuous endogenous variables." (Imbens and Wooldridge, 2009, Lecture Notes 6, p.29)

$$forvol^* = \mathbf{X}\beta + \varepsilon, \quad \varepsilon|\mathbf{X} \sim Normal(0,1) \tag{3.4}$$

where β is $K \times 1$ and $\mathbf{X}$ does not contain a constant. Let $\alpha_1 < \alpha_2 < \alpha_3 < \alpha_4 < \alpha_5$ be unknown **cut points** (or threshold parameters). We define:

$$\begin{aligned}
forvol &= 0 \text{ if } forvol^* \leq \alpha_1 \\
forvol &= 1 \text{ if } \alpha_1 < forvol^* \leq \alpha_2 \\
forvol &= 2 \text{ if } \alpha_2 < forvol^* \leq \alpha_3 \\
forvol &= 3 \text{ if } \alpha_3 < forvol^* \leq \alpha_4 \\
forvol &= 4 \text{ if } \alpha_4 < forvol^* \leq \alpha_5 \\
forvol &= 5 \text{ if } forvol^* > \alpha_5
\end{aligned} \tag{3.5}$$

To derive the conditional distribution of $forvol$ given $\mathbf{X}$, we compute each response probability:

$$\begin{aligned}
P(forvol = 0|\mathbf{X}) &= P(forvol^* \leq \alpha_1|\mathbf{X}) = P(\mathbf{X}\beta + \varepsilon \leq \alpha_1|\mathbf{X}) = \Phi(\alpha_1 - \mathbf{X}\beta) \\
P(forvol = 1|\mathbf{X}) &= P(\alpha_1 < forvol^* \leq \alpha_2|\mathbf{X}) = \Phi(\alpha_2 - \mathbf{X}\beta) - \Phi(\alpha_1 - \mathbf{X}\beta) \\
P(forvol = 2|\mathbf{X}) &= P(\alpha_2 < forvol^* \leq \alpha_3|\mathbf{X}) = \Phi(\alpha_3 - \mathbf{X}\beta) - \Phi(\alpha_2 - \mathbf{X}\beta) \\
P(forvol = 3|\mathbf{X}) &= P(\alpha_3 < forvol^* \leq \alpha_4|\mathbf{X}) = \Phi(\alpha_4 - \mathbf{X}\beta) - \Phi(\alpha_3 - \mathbf{X}\beta) \\
P(forvol = 4|\mathbf{X}) &= P(\alpha_4 < forvol^* \leq \alpha_5|\mathbf{X}) = \Phi(\alpha_5 - \mathbf{X}\beta) - \Phi(\alpha_4 - \mathbf{X}\beta) \\
P(forvol = 5|\mathbf{X}) &= P(forvol^* > \alpha_5|\mathbf{X}) = 1 - \Phi(\alpha_5 - \mathbf{X}\beta)
\end{aligned} \tag{3.6}$$

The model is estimated by maximum likelihood from the log-likelihood function by using Stata. For each observation i, the log-likelihood function is

$$\begin{aligned} l_i(\alpha,\beta) &= 1_{(forvol_i=0)} \cdot log[\Phi(\alpha_1 - \mathbf{X}_i\beta)] + 1_{(forvol_i=1)} \cdot log[\Phi(\alpha_2 - \mathbf{X}_i\beta) \\ &- \Phi(\alpha_1 - \mathbf{X}_i\beta)] + \cdots + 1_{(forvol_i=5)} \cdot log[1 - \Phi(\alpha_5 - \mathbf{X}\beta)] \end{aligned} \tag{3.7}$$

We now amend the above estimation procedure to control for endogeneity of employment status. As suggested by Imbens and Wooldridge (2009), it is convenient to model the source of endogeneity as an omitted variable(s). In our model, these can be confounder latent variables that influence volunteering and are potentially correlated with employment. Motivation and ability are obvious examples. We now disaggregate our vector of regressors, **X**, into **Z**, job, and x_u, where **Z** is a vector of exogenous regressors, job is an endogenous regressor and x_u is the omitted variable(s) that we would like to control for. job is a dummy variable which is equal to one if the respondent reported being employed full-time at the time of the survey. The structural model for the response probabilities, that was specified above, is slightly modified as:

$$P(forvol = j|\mathbf{Z}, job, x_u),\ j = 0, 1, \cdots, 5 \tag{3.8}$$

As with all control function approaches, we need enough exclusion restrictions in **Z** to identify the parameters. We now apply a simple two-step procedure. We first estimate the reduced form for job and obtain the residuals from this regression which we term as $\hat{e}^{CF}$. We plug in these residuals in the second step to control for the endogeneity originating from our inability to directly control for x_u.

Using this approach, we estimate the following reduced form equation, involving probit specification, in the first stage:

$$1_{(job_{it}=1)} = \Phi(\gamma \mathbf{w}_{it}) + \varepsilon_{it} \tag{3.9}$$

where $\Phi(\cdot)$ is the cumulative distribution function for the standard normal, and **w** is a vector of explanatory variables. These include our instrumental variable, $treat$, a dummy variable equal to 1 if the participant is randomly assigned to the Program group. $treat$ is our excluded instrument. For identification, this must not be correlated with the disturbance process of the primary equations and must be highly correlated with the included endogenous regressor. Randomization ensures that the first condition is met, and we test the second condition in first-stage regressions.

As an alternate estimation strategy for the first step, we also estimate a linear probability model (LPM):

$$1_{(job_{it}=1)} = \gamma \mathbf{w}_{it} + \epsilon_{it} \tag{3.10}$$

Our second-stage then becomes:

$$P(forvol = j|\mathbf{Z}, job, \hat{e}^{CF}),\ \ j = 0, 1, \cdots, 5 \tag{3.11}$$

We use a nonparametric bootstrap to compute the standard errors in the second stage[7]. Bootstrap standard errors are obtained by resampling observations (with replacement) from the data in memory 200 times.

Since there is differential attrition in our data between Program and Control groups over time, we conduct a test of attrition bias using the procedure suggested by Becketti, Gould, Lillard and Welch (BGLW 1988). We do not find any statistically significant evidence of attrition bias.

[7]Use of bootstrap standard errors provides consistent standard errors than the usual standard errors, because bootstrapping factors in the uncertainty in predicting the first-stage residuals

It is worth highlighting, especially for the interpretation of our results, that we are essentially estimating the causal effect of the treatment, employment, when compliance is imperfect, so that random assignment generates an instrument for the treatment of interest, employment. As discussed in Angrist and Imbens (1994, 1995), the ratio of the intention-to-treat estimate and the fraction of individuals who were treated in the Program and Control groups can be interpreted as a Local Average Treatment Effect (LATE). In case of partial compliance, this estimate is the average treatment effect for a well-defined group of individuals, namely the group of compliers who are induced by the instrument to take advantage of the treatment. Denoting the outcome of interest (volunteering levels in our case) as Y, the treatment of interest (employment in our case) as T, and the instrument (random assignment) as Z, and recognizing that the control function approach relies on the same kinds of identification conditions as the standard IV methods (Guido and Imbens 2009), our estimate is:

$$\beta = \frac{E[Y_i|Z_i = 1] - E[Y_i|Z_i = 0]}{E[T_i|Z_i = 1] - E[T_i|Z_i = 0]} \tag{3.12}$$

This framework will be helpful for interpretation of our results later because it demonstrates that we are estimating the effect of the treatment (employment) on volunteering for those whose treatment status was affected by the instrument (random assignment). Because the project (CEIP) provides employment to the Program group for 36 months only, therefore the fraction of individuals who were treated (employed) in the Program group[8], is positive in the second wave, but become negative in the next two waves.

[8]No one in the Control group received employment because of the project.

3.6 Estimation Results

We first present our main findings from the second-stage estimation of the effect of employment on formal volunteering, and later describe the results from the first stage.

3.6.1 Second-Stage Results

Table B.6 displays the result from estimation of formal volunteering on employment and other regressors, for Waves-2, 3 and 4 while using residuals from probit estimation in the first-stage. We do not show results for Wave-1 because it is the baseline before participants were randomly assigned to Control and Treatment groups. Thus our instrument, random assignment to treatment, generates a significant change in employment status for participants in Waves 2, 3 and 4 only. Columns numbered Wave 2 CF, Wave 3 CF and Wave 4 CF display results from the control function (CF) approach for Waves-2, 3 and 4 respectively. Columns for Wave 2b, 3b and 4b display corresponding estimation results from Waves 2, 3 and 4, respectively, without controlling for endogeneity. Complete second-stage results from Stata pertaining to columns numbered Wave 2 CF, Wave 3 CF, and Wave 4 CF, respectively, are displayed as Figures B.5, B.6 and B.7. The full output also displays the cutoff points corresponding to the ordinal values taken by our dependent variable ($forvol$). We later also determine the marginal effects of all regressors.

Comparing results for each wave with and without controlling for endogeneity, the most important thing to note is that failure to control for endogeneity severely underestimates the effect of employment on volunteering levels. Thus in Wave-2, whereas using the control function approach generates a coefficient of 0.736 on job which is statistically significant at the 1% level, direct estimation of employment's effect on volunteering generates a coefficient of 0.0475 which is not statistically significant as conventionally calculated. Similarly in Wave-3, the control function approach generates a coefficient of -5.173 on job which is statistically significant at the 1% level, while direct estimation generates a coefficient of -0.170 which is only statistically significant at the 5% level. Similarly in Wave-4, the control function approach generates a coefficient of -4.817 on job which is statistically significant at the 5% level, while direct estimation generates a coefficient of -0.0805 which is not statistically significant at all. These results suggest the importance of instrumenting for employment and of using a control function approach.

The second most important point to note from these results about the effect of employment is that the effect is positive in the first wave (i.e., Wave-2), and is then strongly negative for the next two waves. This is somewhat counter-intuitive at first, but can be explained using the peculiar nature of this project, using the same logic we employed earlier in explaining the local average treatment effect (LATE) framework and that we use later in interpreting the results from the first stage.

Since our identification is coming from the group of compliers, those whose treatment status is affected due to random assignment, the sign flips because the compliance fraction is positive in the Wave-2 survey when the project provides jobs to respondents, and turns negative in the next two waves when project-related employment comes to an end. In all cases, the coefficient on job in our estimation represents the average effect of having a job on levels of formal volunteering for the subset of people who got a job due to random assignment. Thus in Wave-1, the average effect of having a job on volunteering is positive for the group of compliers in Wave-2, and is negative in the next two waves when the compliance ratio becomes negative. Thus while most people have been laid off from jobs that they had obtained due to random assignment, those who are still employed severely cut down on their levels of volunteering. We will be interpreting this further in a later section.

Among other determinants of volunteering, the results from Wave 3 suggest that age has a positive, but declining, influence on levels of volunteering. This is similar to the finding of Menchik and Weisbrod (1987) who suggest that volunteer work seems to follow a life-cycle pattern, increasing with age till a certain age, and falling thereafter. They interpret this finding as tending to support the investment model of volunteering, since the older one is, the shorter the recoupment period of the volunteer work investment and, hence, the less it would be undertaken.

Gender has a significant effect on volunteering level, with males volunteering at lower levels. This is true for all specifications and the effect is statistically significant at the 1% level in all but one specification. This result is borne out by the literature on volunteering. Thus Vaillancourt (1994) reports that men participate significantly less in volunteer work than women. He suggests that this finding indicates perhaps differences between men and women in tastes or in the intra-family allocation of non-market work or leisure time. Menchik and Weisbrod (1987) and Day and Devlin (1996) report a similar finding.

It is interesting to note that wage rate effects are generally positive, but are not statistically significant and prima facie suggest that opportunity cost hypotheses do not seem to carry much support empirically. However, this result can be an artefact of the unique nature of our wage variable, $breswage$ which represents the self-reported reservation wage of the survey respondents at baseline.

Education has a positive, and increasing, effect on volunteering levels that is statistically significant for all levels of education. Thus those who have completed high school tend to volunteer more compared to high-school dropouts. Those who go to college tend to volunteer at higher levels, while university education has even greater positive effect on volunteering levels. Among the different levels of highest education achieved, university education has the biggest effect on volunteering behaviour in terms of magnitude, and this effect is also statistically significant at the 1% level in all three waves. Our findings are in line with earlier studies which suggest that education has a strong, positive effect on volunteering behaviour (Webb and Abzug 2008, Vaillancourt 1994, Day and Devlin 1996). Vaillancourt (1994) interprets this finding to state that career and human capital benefits outweigh the price (wage) effect, while Day and Devlin (1996) interpret this positive relationship as one of the beneficial externalities arising from formal education.

Webb and Abzug (2008) report that individuals in professional, managerial, and military occupations are more likely to volunteer than are individuals in other occupational categories. Unfortunately, we do not have enough variation in our data on this dimension to test this hypothesis.

Marital status does not seem to have a statistically significant effect on volunteering. This is contrary to many earlier studies (Vaillancourt 1994, Menchik and Weisbrod 1987, Day and Devlin 1996) who find significant effects of marital status on volunteering. However, our results suggest that the presence of greater numbers of children in the household increase the frequency of volunteering. A possible explanation is provided by Vaillancourt (1994) who reports that the presence of children aged 0-2 reduces the participation in volunteer work of women while having no significant impact on that of men, while the presence of children aged 3-5 and 6-15 and an increase in their number significantly increase the participation in volunteer work of both men and women. Menchik and Weisbrod (1987) also report that those with children at home appear to do more volunteer work than others, but those with young children volunteer fewer hours than those with older children.

Household income has a small but statistically significant effect on volunteering levels. The positive coefficient on $hhincome$ and the negative coefficient on $hhincomesq$ in waves 3 and 4 suggest that income levels have a positive but declining influence on volunteering levels. Literature generally finds positive income effects (Vaillancourt 1987, Day and Devlin 1996), but Menchik and Weisbrod (1987) report a positive but declining income effect.

Table B.7 displays the result from estimation of formal volunteering on employment and other regressors, for Waves 2, 3 and 4 while using residuals from a Linear Probability Model in the first-stage. This table uses the same reporting pattern of results as in Table B.6, displaying pairs of results for each wave, with the first column for each wave representing estimates from using a control function approach and the second column for each wave representing direct estimates by ignoring the endogeneity issue. Although we get slightly different coefficients, the results exactly match the ones described earlier.

3.6.2 First-Stage Probit Results

Table 3.2 displays the coefficients on random treatment assignment from the first-stage estimation for the four waves of the CEIP project survey data. The complete results from first-stage regressions of employment (job) on the complete set of regressors are displayed in appendix Tables B.8 and B.9 for probit and linear probability models respectively.

Table 3.2: Summary of First Stage Results

	Wave 1	Wave 2	Wave 3	Wave 4
Probit Results	.0477	1.170^{***}	-0.239^{**}	-0.119
	(0.083)	(0.089)	(0.082)	(0.090)
OLS Results	.0099	0.366^{***}	-0.0872^{***}	-0.043
	(0.020)	(0.025)	(0.029)	(0.033)

Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

For full set of regression results, see appendix Tables B.8 and B.9

The top row of Table 3.2 displays coefficients on the random treatment assignment variable ($treat$) using the probit model while the bottom row displays estimation results from the linear probability model. The results match our *ex ante* anticipation. Thus if randomization worked properly, we would expect the coefficient on random assignment to treatment to be insignificant in affecting employment status in Wave-1 corresponding to baseline, that is, before random assignment of participants to the Program and Control groups.

Similarly we would expect our instrument, random treatment assignment, to be significant in affecting employment status in Wave-2 and Wave-3 surveys. The Wave-2 surveys were conducted roughly after 18 months from random assignment, that is, in the middle of the 36-month employment eligibility period for Program group participants. The Wave-3 surveys were conducted roughly after 40 months from random assignment. The employment eligibility period for Program group participants consisted of 36 months, and therefore Wave-3 surveys were conducted after the end of employment eligibility for most Program group participants. However, there was a lag for some participants between random assignment and the beginning of their employment. As a result, some Program group participants were still completing their employment eligibility at the time of the Wave-3 surveys. Wave-4 surveys were conducted 54 months after random assignment, more than a year after the end of employment eligibility for all participants. We would thus expect random treatment assignment to be insignificant in Wave-4.

This is indeed what we find in our estimation. Thus, for probit estimation, the coefficient on $treat$ is statistically insignificant in Wave-1, is statistically significant at the 1% level in Wave-2 and at the 5% level in Wave-3, and is statistically insignificant again in Wave-4. In terms of magnitude, the coefficient is substantial for Wave-2 estimation, and less so for Wave-3 estimation, as one would expect.

In terms of the direction of effect, random treatment assignment has a positive effect on employment in Wave-2 estimation, and a negative effect in Wave-3 estimation. The most intuitive way to interpret this result is to think of the project as comprising two different and opposite interventions. The first intervention consists of offering employment eligibility to participants, and this naturally has a positive effect on the employment outcomes of the participants who had been randomly assigned to the Program group. The second, and opposite intervention, comprises the end of employment eligibility for these participants, when their employment with the project ended. This naturally has a negative effect on their employment status. The unique nature of these interventions allows us to explore the effect of having a job on volunteering under different environments, namely when most of the compliers are employed and later when most of the compliers are not employed. Unlike previous studies, this also focuses our attention on the precise channels through which employment has an effect on volunteering.

3.6.3 Effect of Employment on Informal Volunteering

For comparison, we also examine whether employment has any effect on the level of *informal* volunteering, which entails providing unpaid help to friends or neighbours, and which could take the form of providing support to the sick or elderly, doing house or maintenance work, driving people to appointments or baby-sitting for someone who is not a relative or a member of their household.

Like formal volunteering, the responses on frequency of informal volunteering during the past 12 months were reported on an ordered response scale that takes six values - 0 (never), 1 (less than once a month), 2 (once a month), 3 (once a week), 4 (few times a week), and 5 (everyday). A similar approach was used in employing an ordered probit model for estimating the effect of employment on levels of informal volunteering, using random assignment to the treatment group as an instrument for employment, and employing the control function approach.

Table B.10 displays the results of such estimation which uses probit estimation in the first-stage. Columns for Wave 2 CF, Wave 3 CF and Wave 4 CF display results from the control function approach for Waves 2, 3 and 4, respectively. Columns numbered Wave 2b, 3b and 4b display corresponding estimation results from Waves-2, 3 and 4, respectively, without controlling for endogeneity.

The results show that employment status does play a significant role in determination of levels of informal frequency in some waves. In both waves 2 and 3, being employed generates a decrease in levels of informal volunteering in the empirical specifications that do not involve a control function approach. This decrease is statistically significant at the 5 percent level (as conventionally calculated). Employment status is not statistically significant in any specification in Wave-4. It appears that employment status is not endogenous in the determination of levels of informal volunteering, and that the informal employment effect is smaller than the formal employment effect.

Among other regressors, gender plays a strong role in Wave 2 only which is positive and statistically significant at the 1% level. Comparing these results with the ones for formal volunteering, gender (being male) has a positive effect on levels of informal volunteering whereas it had a negative effect on levels of formal volunteering. Finally, educational status and having more kids is associated with a positive effect on levels of informal volunteering. These results are similar to the ones for formal volunteering.

3.7 Interpretation of Results

Our results suggest that age, gender, educational levels, and marital status are major predictors of volunteering behaviour. This finding, and the direction of our effects, are broadly in line with earlier studies on the determinants of volunteering behaviour (Menchik and Weisbrod 1987, Vaillancourt 1994, Day and Devlin 1996, Webb and Abzug 2008). However, our results on the effect of employment on volunteering behaviour are different from some existing findings, primarily because these reflect a change in the direction of the effect between different waves. This interesting finding has been made possible both because of the unique nature of our experimental interventions and because of the panel nature of the data that we analyse.

Although we are currently exploring this finding further, it is possible to offer a plausible preliminary explanation. The positive effect of employment on volunteering in Wave-2 is associated with a context where employers have an incentive in hiring the project participants, and where the organizational settings where these participants are hired provide significant opportunities, and perhaps some encouragement, for participants to engage in formal volunteering. The negative effect of employment on volunteering in the next two waves comes in the context of employment eligibility for these participants and when they no longer face the same organizational environment. It is possible that those who continue to be employed after the end of project employment are at the higher end of the skill (and unobserved ability) distribution of the participants, and that they either cut back on their volunteering levels in an environment where jobs are scarce, or were already volunteering at lower levels and were focusing more on their jobs.

It will be interesting to explore these hypotheses in further work, especially in view of some of the previous literature on volunteering as a “conscience good”, and some of the recent literature on relational volunteering. Thus Freeman (1997) suggests that standard labor supply explanations of volunteering account for only a minor part of volunteer behavior. Our results suggest that organizational context may be an important determinant of the levels of formal volunteering, and thus our results focus attention on the precise channels through which employment and volunteering may be connected. This interpretation is similar to the one offered by Freeman (1997) who shows that volunteers have high skills and high opportunity costs of time, and that many volunteer only when requested to do so. He suggests that volunteering is a "conscience good or activity" - something that people feel morally obliged to do when asked, but which they would just as soon let someone else do.

Freeman argues that two factors underlie the response of individuals to requests to volunteer: first, that people value the particular charitable activity; second, that the request carries some "social" pressure with it. He adduces some empirical evidence on the importance of being asked, and claims that this interpretation is also suggested by other evidence (Independent Sector 1986, Rockefeller Brothers Fund 1986, Freeman 1993, Thomas and Finch 1990). A 2000 survey of Illinois residents found that 38 percent of the employees worked at companies that sponsored a community project for which they were encouraged to volunteer; 28 percent worked at companies that gave incentives or recognition to employees who volunteer; and a quarter worked at companies that gave money to the organizations for which they volunteered (Profile of Illinois 2001).

We explore the channels connecting employment and volunteering from our data. Table B.11 shows the responses of participants when asked about the reasons for engaging in formal volunteering. It is important to note that 85.9 percent cite 'enjoy helping other people' while 15.7 percent cite 'to help cause in which they personally believe' as their reasons. Job-related reasons are reported with much less frequency: thus only 1.2 percent mention improving job skills, 1.7 percent mention improving job opportunities, 0.8 percent mention required by school/employer/government, and 1.4 percent mention 'already work for volunteer organization'.

However the survey only asked people about the reasons for volunteering[9], and did not explore if the respondent was asked to volunteer. Self-reported reasons for volunteering may also not be completely reliable. It is possible that respondents who volunteered for job-related reasons or because of social sanctions may have internalized the more socially acceptable reasons for volunteering.

CEIP jobs were primarily in the "third" sector, which brought project participants into contact with non-profit organizations, which have community-driven missions that historically depend on volunteers in their day-to-day operation. On the other hand, CEIP required active program group members to work 35 hours per week, which may impose serious time constraints for some participants who could prefer to use their non-CEIP time for pursuits other than volunteering. It is not obvious *ex ante* as to which of these arguments is more important.

When we look at the specific channels through which respondents volunteer, it appears that job-related reasons are significant. Table B.12 displays the forms in which the respondents exercised formal volunteering for Waves 2, 3 and 4. Thus the most important forms of volunteering are organizing or supervising activities for an organization; canvassing, campaigning or fundraising; and collecting, serving or delivering food as a volunteer through an organization. All of these forms of volunteering appear to be strongly related to the employment context. The least important forms of volunteering are providing information to influence public opinion; providing care or support; and volunteer driving on behalf

[9]This question was asked only in the Wave-2 survey.

of an organization. These forms of volunteering appear to be less closely related to the job-related reasons. Interestingly, the most significant variation is experience in the forms of volunteering that appear to be related to jobs to a greater extent. This suggests that not only the levels of formal volunteering but also the forms it takes may be closely related to the specific employment context of participants. In ongoing work, we hope to explore this interpretation more formally.

3.8 Conclusion

This study uses data from the Cape Breton Community Employment Innovation Project (CEIP) to estimate how changes in employment affect volunteering behaviour. Because of potential endogeneity issues, this study uses a consistent estimation strategy based on a control function approach. The results provide strong evidence that employment status has a substantial effect on formal volunteering behaviour, but a weaker effect on informal volunteering. However, the results also suggest that this impact is context-dependent and not uniform. Furthermore, the effect of employment on formal volunteering appears to be mediated by the precise channels through which the two are connected.

Our results may need to be qualified by the fact that these are not based on representative samples, but on rather low-income samples. Another possible qualification comes from the evidence that suggests that Atlantic Canada, where Cape Breton, is situated, tends to be the most generous region in Canada in terms of charitable donations and volunteering (Hall et al. 2009).

The findings of this study have significant policy implications. Thus, these suggest that working for pay does not necessarily substitute working without pay, or volunteering. This further suggests that policy interventions aimed at generating employment opportunities can have spillover effects in terms of encouraging volunteering behaviour that, in turn, may generate positive social and economic outcomes. But our findings also suggest that for policy interventions targeting employment outcomes to also affect non-labour market outcomes, the channels through which the two are connected need to be factored in.

In ongoing work, we are exploring the channels through which employment has an effect on volunteering behaviour. We are also currently exploring amending our theoretical specification in order to treat volunteering also as an investment good. The ongoing work involves examining a dynamic specification of volunteering behaviour by taking advantage of the panel nature of the data. Further possible work will involve estimating the equation for time spent on volunteering jointly with one for time spent on market labour.

An important extension of this work is to combine the analysis in this paper with the one in the previous paper. This is the direction suggested by Apinunmahakul and Devlin (2008) who explore the interaction of social networks with private philanthropy.

Chapter 4

Impact of Decentralization on Provision of Public Services: Evidence from Pakistan

4.1 Introduction

Over the past three decades, decentralization of political, administrative, and fiscal authority to elected local governments has become a significant part of the institutional reform initiatives in many countries, especially in the developing world. Decentralization reforms have been aimed at improving service delivery by providing greater control rights to citizens, and thereby at increasing the accountability of service providers to local communities (Bardhan and Mookherjee 2006, World Bank 2003).

However, the actual impact of decentralization on the responsiveness and accountability of public services is still an open empirical question. This is on account of the fact that it is hard to come up with a consistent estimation strategy that not only compares pre- and post-decentralization public provision but that also takes account of unobserved effects that may confound this estimation. Because fiscal and administrative decentralization often lags electoral decentralization, this makes it hard to attribute changes in post-reform expenditures to local political structures. Finally, we may need to go beyond just estimating impact since we have limited understanding of the political channels through which an impact can come.

This study estimates the impact of decentralization on the provision of public services. We use empirical evidence from Pakistan for this study because it has undergone a major process of political decentralization. This involved decentralization of key public services and resources to the local levels, establishment of new political structures for local governments, and transfer of authority for provision of decentralized public services to local elected politicians. These changes have resulted in conferring authority for making expenditure allocations to elected politicians and in making local bureaucracy accountable to local politicians. This presents us with a natural experiment, which we use in order to identify the impact of decentralization empirically. To be specific, we exploit the fact that not all sectors have been decentralized and within each sector, not all activities have been decentralized.

4.1.1 Research Question, Methodology and Potential Contribution

This study seeks insights into the relationship between decentralization and service delivery by examining the expenditure patterns of local governments to gauge their sectoral priorities. It primarily examines if decentralization has an impact on the provision of public services, and attempts to identify the political channels through which this impact arises[1].

Traditional approaches to measuring the empirics of decentralization, say through a before-after methodology, are open to the challenge of not addressing some of the concurrent changes taking place along with decentralization. Since these studies examine variations in outcomes across sectors, this can confound estimation since other trends are likely to reflect cross-sectoral variation. For instance, a general progressive trend may mean spending more on social sectors in addition to decentralization - neither causes the other but both are related.

This study makes a number of important contributions. Specifically, it attempts to provide some evidence regarding whether decentralization impacts public allocation decisions. It improves identification using a difference-in-difference (D-D), and later a difference-in-difference-in-difference (D-D-D) approach, but one based on utilizing institutional and contextual knowledge to produce consistent results. It exploits *within* sector differences, and looks at what public goods are affected by decentralization and why. The unique design of Pakistan's far ranging recent local government reforms allows us to attribute changes in post-reform expenditures to local government political decision-making.

A caveat on the results in this study may be pointed out at the outset. This study deals with budgetary outcomes and not real socioeconomic outcomes. This approach, however, is in line with most studies on public good provision that tend to use expenditure figures rather than data on actual provision. See, for example, Alesina et. al. (1999) and Faguet (2004). In future work we would hope to examine socioeconomic outcomes.

This chapter is organized as follows: we discuss the literature on decentralization in the next section, and then describe the context of decentralization reforms in Pakistan. We

[1] In on-going work, we are exploring an additional source of variation, the major social sector aid infusion to Pakistan after the events of September 11, 2001 (9/11), in order to better identify these channels.

then discuss our data and identification strategy, and present the results for D-D and D-D-D approaches. Finally, we identify the political channels that seem to be driving our results.

4.2 Literature

Two arguments are commonly given in favor of decentralization. One, the preference-matching argument, hypothesizes that decentralization improves allocative efficiency by allowing greater differentiation in the provision of public goods and services. This reflects the belief that, because local governments are closer to the people than the central government, they are better informed about the preferences and circumstances of the residents.

The second argument posits that decentralization increases the accountability of government. Proponents of this argument contend that people tend to be more aware of the actions of local governments than they are of the actions of the higher levels of government because local governments are closer to their constituents (Shah 2006). These arguments can be traced to Stigler's (1957) two principles of jurisdictional design: (1) the closer a representative government is to the people, the better it works, and (2) people should be able to vote for the kind and amount of public services they want. Oates (1999) provides a more formal expression of this argument. Important papers on the political accountability literature include Besley and Coate (2003), Persson and Tabelini (2000), and Seabright (1996).

As is common in the economics literature, we look at accountability and responsiveness in terms of induced outcomes. According to Przeworski et al. (1999): "Governments are 'accountable' if citizens can discern representative from unrepresentative governments and can sanction them appropriately, retaining in office those who perform well and ousting from office those who do not." The electoral accountability mechanism works through governments getting sanctioned at elections on the basis of their performance in office. Thus, in order to survive in office, governments tend to perform in a way that satisfies the majority of voters.

Bardhan and Mookherjee (2006) argue that the competitive pressure associated with winning local elections can foster greater political accountability among governments. Electoral accountability can also work through “yardstick competition” (Besley and Case 1995) where voters look at public services and taxes in other jurisdictions to help judge if their government is wasting resources, through inefficiency or rent seeking, and deserves to be voted out of office. Myerson (2006, 2009) shows that decentralization can promote national democracy by increasing democratic competition as it provides greater opportunities for local elected leaders to build their reputations enabling them to challenge the incumbents at the centre.

However, Bardhan and Mookherjee (2000) also highlight the risks associated with decentralization in terms of possible ‘elite capture’. Seabright (1996) shows that the effect of decentralization on accountability can go either way. This is because, while the value of elected office is higher at nationwide elections than at local elections, local elections provide sharper incentives to politicians to perform well as they are elected on the basis of their local instead of average competence.

Some authors have studied the role of political market imperfections in service delivery and economic development and found that policy breakdowns and the inability of politicians to make credible pre-electoral promises to voters leads them to provide direct targeted goods at the expense of broad universal public services (Keefer 2004, 2005; Keefer and Khemani 2003).

Some authors have argued that decentralization can mitigate two obstacles to efficient public sector delivery - divergence in public good preferences among large groups of citizens, and difficulties concerning large group of citizens seeking to hold public officials accountable for their performance (Azfar et al. 2001; Keefer et al. 2006).

Hindricks and Lockwood (2005) suggest that political accountability may be problematic when elected representatives either have policy priorities different from those of the electorate or are subject to lobbying by interest groups or to political clientelism. Political clientelism refers to the proffering of material benefits in return for electoral support and can take the form of patronage and vote buying (Stokes 2007).

4.2.1 Empirical Evidence on Decentralization

The evidence on the impact of decentralization on service delivery is mixed. Some studies, such as Faguet (2004) on Bolivia, demonstrate that decentralization resulted in a substantial shift in public spending in favor of smaller and poorer municipalities. Bardhan and Mookherjee (2004), show that inter-village allocations of credit, resources for local infrastructure and employment for the poor, and development grants from upper levels of government in the Indian state of West Bengal exhibited poorer targeting than allocations of these resources within villages by local governments.

Other studies argue that decentralization increased elite capture and corruption. A World Bank study of Indonesian village governments established in 1979 showed that accountability of village heads to the villagers was very weak, with a negligible number of village proposals included in district budgets (World Bank 2001). A cross-country investigation by Triesman (2002) found that decentralization was significantly associated with measures of corruption. There is also some evidence on a negative correlation between ethnolinguistic polarization at the local level and access to public goods and development outcomes (Alesina, Baqir and Easterly 1999).

Ahmad et al. (2004) explain this mixed record by hypothesizing that improvements in public service delivery require three strong relationships of accountability between the different actors in the service delivery chain. First, citizens must be able to hold policy-makers accountable; second, policy-makers must be able to hold service providers accountable; and third, the inter-governmental framework between national and local policy-makers must be conducive to improving service delivery. The authors refer to these relationships as the 'long-route of accountability' as opposed to the short route under which in a private, competitive market poor people, as customers, could hold providers directly accountable. Weaknesses in public service delivery can be attributed to breakdowns in any one or all of these links.

Recent studies conducted by political scientists and economists (Crook and Manor 1998, Tendler 1997, Bardhan and Mookherjee 2006, Chaudhuri 2006, Besley et al. 2003, Litvack et al. 1999) show that the success of decentralization in terms of accountability and in improving allocative efficiency is conditional upon both the design of decentralization and the political dynamics of competition and socio-economic structures. Implementation of new political, administrative, and fiscal measures also has a bearing on their outcome (Crook and Sverrisson 2001; Guess 2005).

4.3 Context of Decentralization in Pakistan

Pakistan has had a local government tradition since colonial times. Since independence, two further attempts to devolve power in Pakistan have been made. These include the Basic Democracies Order in 1959 under President Ayub Khan and the 1979 Local Government Ordinance of General Zia ul Haq. Previously, the local government system did not have a meaningful role because local governments were rather inactive and most of the government functions were carried out at the provincial level. However, both sets of reforms were dismantled by subsequent governments. In our earlier work (Cheema, Khwaja and Qadir 2006[2]), we provide an historical overview of decentralization in Pakistan from the pre-independence period to the current system.

The devolution plan introduced by President Musharraf's government in 2000 represents the most comprehensive and far-reaching initiative so far in transforming the local governments[3]. Devolution significantly changed public service delivery by passing the authority to allocate expenditures for various services from the provincial to the district level and by introducing political and administrative accountability mechanisms for elected local governments. Further, the Musharraf plan decentralized the determination and enforcement of laws related to property and labor rights as well as access to justice.

These structural changes introduced under decentralization aim at influencing the incentives within the public sector through a mixture of political, fiscal, and administrative measures that both empower local governments and make them accountable for public service delivery (Manning et al. 2003). Two types of incentives to improve service delivery under devolution were created. First, it was assumed that citizen power, in the form of 'voice', can create incentives for local governments to improve public service delivery by allocating resources efficiently and holding providers accountable for service delivery outcomes. Second, devolution created a mechanism that ensures that local governments have authority over front-line service providers.

[2]I was a coauthor on this earlier work.

[3]The reforms were first introduced through the Devolution of Power Plan, 2000. These were subsequently promulgated through the Local Government Ordinance, 2001. Local Government elections were held during 2001 and these local governments were formally established on August 14, 2001.

4.3.1 Overview of Reforms

The decentralization reforms replaced the deconcentrated[4] government structure that was previously in place with a new one that includes three levels of elected local government: union councils (lower tier), tehsil councils (medium tier), and district governments (upper tier). Overall, there are 96 district councils, 337 tehsil councils, and 6,022 union councils in Pakistan, and 34 district governments, 122 tehsils, and 3,464 union councils in Punjab province, the focus of this study. Each level of local government has elected councils, nazims (mayors), and naib (deputy) nazims. Decision-making power and control over the financial resources largely reside at the district level. This study examines the budgetary allocations of district governments only. Pre- and post-decentralization structures are shown in appendix Figures C.1 and C.2.

The reforms established intergovernmental political linkages by ensuring that two-thirds of the members of the tehsil and district councils are the elected nazims and naib nazims of union councils. The remaining one-third of the seats in each council are reserved for women, peasants, and minorities, who are elected indirectly by the directly elected union council members. The heads of district councils (nazims and naib nazims) are elected indirectly by union council members on a joint ticket. The union council is composed of 21 directly elected members[5], include the union nazim, naib nazim, councillors and reserved seats for women, minorities, labourers and peasants.

[4]Deconcentration is a term referring to the transfer of administrative responsibility for specified functions to lower levels within the provincial/federal government bureaucracy on spatial basis. In our context, it refers to the pre-decentralization system of provincial bureaucracy functioning at the district level that was accountable to the provincial authorities.

[5]The 2005 amendments to the Local Government Ordinance have reduced union council membership to 13.

4.3.2 Decentralization of Service Delivery Functions

Under the reforms, the scope of local governments has been considerably increased by decentralizing key provincial functions to the district and lower levels. Most significantly, budgeting, planning and development functions that were previously performed by provincial bureaucrats have been transferred to the district and lower levels. For example, primary and secondary education, water and sanitation, roads, transport, agriculture and primary healthcare have been assigned to local governments. For a list of decentralized and non-decentralized sectors, see Table C.1. These changes have translated into a considerable increase in the districts' share of consolidated provincial and local government expenditure (Cheema et al. 2006 and Manning et al. 2003). Protagonists of the reform have expected that the decentralized expenditure assignment will strengthen the accountability of governments to citizens by reducing the distance between citizens and service providers, thereby allowing citizens to better monitor the working of government.

4.3.3 Engendering Electoral Accountability

Provincial service delivery departments at the district and local levels have been placed under the authority of elected governments at these levels. This has significantly empowered the local elected tier and created a new form of accountability for the provincial bureaucracy. Thus local bureaucracy, which was previously accountable to un-elected provincial bureaucrats, is now accountable to an elected representative of the citizens. In particular, district bureaucrats who were previously accountable to unelected provincial bureaucrats now have to work under an elected district mayor. Again, this change was expected to increase the responsiveness of administration to citizenry.[6]

4.3.4 Decentralization and Public Service Delivery in Pakistan

Recent literature has shown that Pakistan's sixty-year development experience has resulted in a significant social development gap relative to comparable economies. The literature identifies that Pakistan's poverty of social outcomes in part reflects a bias in public sector expenditures against the social sectors. Pakistan's poor social development outcomes have been attributed to its governance structures that weakened political and bureaucratic accountability to citizens and resulted in political economy and governance failures (Keefer 2002; Keefer et al. 2006). Easterly (2003) has described Pakistan's development experience as the paradox of 'growth without development'. The 2001 local government reforms are an attempt by the Pakistani state to rectify these failures by redesigning political, electoral and administrative structures and relationships at the local level.

The literature identifies three key features for these accountability failures. First, most decisions regarding service delivery, including budgetary allocations, were centralized and concentrated in the hands of the provincial and federal governments. This centralization resulted in a lack of accountability because it created a large distance between citizens and key decision makers, thereby weakening the ability of the former to monitor the actions of the latter.

Second, the bulk of services were delivered through the deconcentrated provincial bureaucracy that was accountable to the higher tier provincial bureaucracy and not to citizens.

[6]The pre-decentralization head of district administration, the Deputy Commissioner (D.C.), who used to report to the provincial bureaucracy, has now been replaced by the District Coordination Officer (D.C.O.), who reports to the district nazim. The D.C.O. in turn supervises a team of officers who head each of the decentralized departments at the district level. In Punjab, administrative authority - i.e., authority over appointment, promotion, transfer, and disciplinary proceedings - over non-officer staff in the decentralized departments has been transferred to the districts.

This created a disjuncture between service providers and citizens, weakening the accountability of the former to the latter.

Third, the political system in Pakistan has been the domain of historically entrenched interests with powerful politicians acting as patrons to selected local level clients because of unequal control and ownership in economic, political and social domains (Hussain 1999, Easterly 2003, and Gazdar 2000). In this view, political bargaining at the local level was confined between the local un-elected patrons and the higher-tier elected politicians and/or the deconcentrated provincial bureaucracy. An important cause of accountability failures was the un-elected nature of local level patrons who influenced decision-making and could control the local level without being accountable to a wider citizenry.

Decentralization in Pakistan has attempted to target each one of these accountability failures. It gives political executives at the district level considerable 'authority' and 'autonomy' over the allocation of district development schemes that provide local public goods. There is some evidence to suggest that the direct accessibility of local policy-makers to citizens is considerably greater than that of provincial and national policy-makers, and that local government elections, especially in Punjab, are as competitive as provincial and national elections in terms of voter turnout and number of contesting candidates (Manning et al. 2003, Hussain 2008). The combination of electoral decentralization with the increased authority of nazims over local public good scheme allocations allows us to attribute post-reform provision changes to local political decision making.

The actual outcome of the decentralization may, however, work differently from its theoretical effects for a number of reasons. First, the decentralized system is not complete (Ajmal and Bari 2006; Birner et al. 2006; Zaidi 2005). Second, indirect elections of district nazims may limit their incentives for a more efficient provision of public services.

The district nazim is granted a large amount of decision-making power with considerable control over resource allocation under the decentralized system, but an indirect election introduces weaker electoral constraints and may weaken the ability of voters to hold them accountable.

The success of decentralization also depends on the conditions for political competition and accountability being more favorable at the local level and on the public being able to demand the efficient provision of public services. Keefer et al. (2006) suggest that the policy horizons of local officials may be more uncertain than those of federal and provincial officials because of the past history of decentralization in Pakistan. These uncertain policy horizons may encourage local politicians to engage in rent seeking rather than focusing on service provision.

Another key issue is that context matters in terms of the political mechanism for allocating public resources. Thus provincial governments have not been keen on local governments due to a conflict of interest between provincial and local politicians. Although there exists a bureaucratic hierarchy in which the district bureaucracy reports to the provincial bureaucracy, there is no parallel political hierarchy. Thus district politicians are not only competing amongst themselves, but also against provincial politicians. This is important in our context since it matters as to whether it is the district politician or the provincial politician who is allocating resources.

This overview of decentralization reforms in Pakistan leads us to form certain prior expectations with regard to the impact of such reforms. Thus, the massive and sweeping nature of the reforms and the transfer of authority for provision of decentralized public services from bureaucrats to local elected politicians lead us to expect substantial changes in post-decentralization levels and composition of public provision, unless the preferences of bureaucrats and politicians do not differ much. Secondly, given the history of clientilist politics and given the imperfections in the decentralized political arrangements (non-party based, indirect elections for nazims, etc.), one would expect that the distortions in public good provision pointed out by earlier studies will not be entirely mitigated. For instance, Keefer et al. (2006) report for the pre-decentralization period evidence from more than one hundred villages that: "the competition for votes leads to school construction but not improvements in school quality". Thirdly, one would expect local politicians to allocate public goods in response to their electoral concerns, although it is not ex ante obvious if electoral competition provide enough discipline to politicians to adopt welfare-maximizing policies.

4.4 Data and Identification Strategy

4.4.1 Data Set

The analysis of this study relies on a unique data set from Punjab province of Pakistan. We focus on Punjab as district governments there have been vested with significant fiscal discretion in terms of setting expenditure priorities. It has decentralized forty percent of the expenditure assignment to local governments (Manning et al. 2003). It also constitutes about 60 per cent of Pakistan's population (approximately 90 million people). Our data set comprises budgetary allocations for all sectors, adjusted for inflation[7], from 34 districts of Punjab and covers the period before and after decentralization. The base fiscal year for the data is 2001-02 and comprises our benchmark for allocations before decentralization since allocations were decided by the provincial government for this year. For the fiscal

[7] The estimation results are robust to estimation in nominal or real terms, but we only display the results for estimation in real terms.

years 2002-3 onwards, district budgets were prepared by district governments. Our data covers the post-decentralization fiscal years 2002-03, 2003-04 and 2004-05, but the current version contains estimation results up to the year 2003-04.

It is worth emphasizing that compilation of this data set constituted a major part of this study. Although all constituent parts of this data set comprise public information, these provincial and district budgets have not been compiled previously[8]. Such compilation on a uniform standard has also proved to be a daunting exercise. Decentralization practically involved establishing a third tier of government at the district level. The process involved developing the capacity of district governments for producing budgetary documents, a process that took a few years to complete. Development and adoption of uniform/standardized rules and chart of accounts also took a number of years: a new accounting model was only established across board from the fiscal year 2005-06 onwards. As a consequence, the district budgets for the first four years, comprising our sample, were prepared on different formats/templates. This made the job of creating a uniform and consistent database quite difficult. The provincial budgets also had to be disaggregated at the district level to make these comparable to district budgets. It thus took more than two years to compile this data set. Components of our primary data are publicly available and in theory, a reader can access the data, aggregate it and replicate our results.

The empirical analysis of this study focuses on expenditure allocations because district governments continue to have limited revenue bases and are highly dependent on provincial and federal budgetary transfers. However, these intergovernmental budgetary transfers are non-discretionary in character and are determined by a ‘rule-based’ fiscal resource transfer system through the Provincial Finance Commission. These transfers are no longer lapsable

[8]Interestingly, even the government (Finance Ministry) does not compile district budgets in hard or soft format.

and continue to be retained by the respective local governments.

For identification of political channels, we rely on additional sources of data. These include the Household and Population Census of 1998, and Local Government election results for the years 2001 and 2005. Data from the 1998 Census, conducted much before decentralization reforms had been contemplated or introduced, serve as controls in our empirical specification for political channels. The first Local Government elections under the decentralized system were held in the year 2001, and following the completion of the elected members' four-year tenure subsequent elections were held in the year 2005. We have access to data on the complete election results for the district nazim for these two elections. We use the election results and expenditure allocations to estimate the probability of reelection of the district nazim as a function of the expenditure allocation in various sectors.

4.4.2 Classification of Expenditure Allocations

Expenditure allocations fall into two major categories: development, and current. Development expenditure reflects investment, mostly capital works, designed to keep intact, enlarge and improve the physical resources of the province. The most common form of development expenditure constitutes expenditure on the establishment of new facilities and infrastructure.

Current expenditures relate to the costs of running current services and repairing and maintaining existing facilities. These include both wages and operational costs. Current establishment expenses cover expenditures on the permanently sanctioned establishment (employees) and include salaries and related expenses including allowances.

Current non-establishment expenses, also called non-salary operational expenses, comprise all expenditures that are neither salaries or expenses for building new facilities nor infrastructure. These specifically comprise operational expenses for maintaining and running existing public facilities, services and programs that are classified further as repair and maintenance (R & M) and commodities and services (C & S) costs.

That establishment expenditures act as a control is crucial to our identification. This holds true because public sector employees of decentralized sectors continue to remain provincial employees even after the introduction of decentralization and their wages are determined de facto by the higher (federal and provincial) governments[9]. Thus, although nominally the establishment expenses are passed on by the province to the districts, in reality the districts merely act as a clearing mechanism in passing on these salaries to the provincial employees. The relevant law, Local Government Ordinance 2001, in fact prevents district governments from hiring any new employees, or firing existing ones. This lack of discretion to district governments in hiring or firing employees or changing their salaries justifies our use of establishment expenses as a natural control group in our specification.

To recapitulate, the district budget comprises the following allocations for each sector:

1. Current establishment expenses,

2. Current non-establishment expenses,

3. Development expenditures.

[9]Salary levels for the provincial employees are adjusted for inflation from time to time by the province, but in reality, the province follows the adjustments made by the federal government. We control for these changes.

Summary statistics are given in Table C.2 which shows aggregate expenditures of district governments, in millions of nominal Pakistani rupees, for different types of expenditures - establishment, non-establishment and development. Each observation in our data set represents expenditure on a specific sub-sector (e.g., health administration, vaccination, outreach facilities, etc.) within each sector (health, education, etc.) of a particular type (establishment, non-establishment, development) for a particular district for a given year[10]. Thus establishment expenditures in nominal terms increased from Rupees 29.23 billion in the base year to Rupees 36.58 billion in 2002-03 and Rupees 44.53 billion in 2003-04. Non-establishment expenditures increased from from Rupees 6.36 billion in the base year to Rupees 7.61 billion in 2002-03 and Rupees 6.99 billion in 2003-04, while development expenditures increased from Rupees 11.14 billion in the base year to Rupees 11.2 billion in 2002-03 and Rupees 18.19 billion in 2003-04.

Tables C.3 and C.4 display summary statistics for aggregate district expenditures separately for twelve sectors, in millions of nominal Pakistani rupees. It is obvious from these tables that there exists considerable variation in the size and expenditure patterns across sectors. Thus Education and Health are the largest sectors, while Housing and Public Health are the smallest sectors. Table C.5 displays aggregate expenditures at the provincial level for four non-decentralized sectors. It is worth emphasizing that whereas Table C.2 presents district expenditures aggregated over the entire range of decentralized sectors[11], the Table C.5 presents expenditures for only a small subset of the non-decentralized sectors. Thus these sectoral expenditures at the provincial level, when aggregated, do not comprise

[10]Some sectors, like Education, have a number of sub-sectors. Many observations are of a unique expenditure type , and do not have matching expenditures on corresponding expenditure types for the same sub-sector. For instance, many observations represent development expenditures for specific projects that do not have corresponding establishment and non-establishment expenditures for the same sub-sector.

[11]We, however, present expenditures for only 12 of the 17 decentralized sectors in the next two tables, these being the more interesting sectors

the entire universe of provincial budget which also includes several other sectors.

4.4.3 Identification Strategy

Our empirical strategy attempts to uncover the political preferences of local policy-makers by analyzing the sectoral expenditure priorities of district governments. We exploit the fact that local policy-makers have significant discretion over the funds at their disposal for certain types of expenditures.

Our identification strategy relies on the following:

1. Not all sectors are decentralized. For instance, school education (primary and secondary) is decentralized, but college education is not.

2. Within a given sector, not all activities are decentralized. Thus development and non-establishment expenditures have been decentralized, while establishment expenditures have not been decentralized.

We exploit, as part of our identification strategy, the following differences in expenditures:

- over time (before-after)

- across type of expenditures - establishment (**Control**) versus development and non-establishment (**Treatments**)

- across sectors based on the decentralization principle (decentralized versus non-decentralized sectors)

Our "treatment" groups are non-establishment (operational) expenditures in decentralized sectors which are less discretionary, and development expenditures which are almost fully discretionary at the local level. Thus, instead of comparing across different types of activities (for instance, education versus police), our primary source of identification comes from comparing across different types of activities within a given department. Thus, to the extent that there are differences across sectors that may change over time, this does not confound our analysis. However, we later make use of these differences to provide an additional, third source of, difference.

Our difference-in-difference (D-D) estimator compares outcomes before and after a policy change for a group affected by the change (Treatment Group) to a group not affected by the change (Control Group)[12]. Early examples of the use of this approach include Card and Krueger (1992).

Our identification strategy satisfies the key assumptions behind difference-in-difference estimation. Thus our control and treatment groups have common support since both pertain to the same sectors and the same territorial jurisdictions. The parallel trends assumption is hard to test for the pre-decentralization period because disaggregated budgetary data only exist for one year prior to decentralization. However, the assumption seems to hold well for aggregated provincial data prior to decentralization.

There is a recent literature that highlights the limitations of D-D estimators. In an important paper, Bertrand, Duflo and Mullainathan (2004) demonstrate that D-D standard errors underestimate the sampling standard deviation of estimators because of a possibly severe serial correlation in the data series. Nonparametric techniques and block bootstrapping have been known to perform well. Our identification is not likely to suffer from this problem, and thus bootstrapping does not make an appreciable difference both because the

[12]If T_B, T_A are treatment group outcomes before and after the treatment, and C_B and C_A are control group outcomes before and after the treatment, then $D - D = (T_A - T_B) - (C_A - C_B)$. In regression terms: $Y = \beta_0 + \beta_1 TRT + \beta_2 POST + \beta_3 TRT.POST + \varepsilon$, where TRT is a dummy for treatment group, and $POST$ is a dummy for the post-intervention period, then $D - D = [(\beta_0 + \beta_1 + \beta_2 + \beta_3) - (\beta_0 + \beta_1)] - [(\beta_0 + \beta_2) - (\beta_0)] = \beta_3$.

number of periods in our estimation is small (three in the current version), and because of less likelihood of group error effects since our unit of observations corresponds to the unit of variation.

Our context and identification strategy justify the view that our results are broadly representative. The institutional context is similar to the one existing in many other countries, although there may be more variation on ethnic and other dimensions in other contexts. The events of 9/11 do not directly affect our identification since the decentralization experiment was implemented prior to these events. 9/11 does have some indirect effects which we are exploiting in our ongoing work. The post 9/11 years, comprising our study period, witnessed rapid economic growth in Pakistan, partly owing to the external aid infusion. This did not affect decentralization directly but the resulting increase in revenue collection at the national level did result in large increases in transfers to the provinces which passed a proportion of it on to the local governments. These transfers had no strings attached because these were channeled through the formula-based, non-discretionary process of the Provincial Finance Commission. In order to control for this revenue effect, we are currently estimating the impact in terms of budgetary shares and are exploring quantile regressions.

The events of 9/11 did have an indirect influence on decentralization in Pakistan, because a significant part of the resulting external aid infusion was exclusively targeted to the higher governments (federal and provincial). A substantial part of the funding went to the social sectors which was spent by the provincial governments in districts through vertical, performance-based grants that were subject to conditionalities on local government performance. In ongoing work, we are exploiting 9/11 as an additional source of exogenous variation.

4.5 Estimation Results

This section presents the main findings of the empirical analysis. Our basic specification is:

$$\begin{aligned} lnexp &= \beta_0 + \beta_1 post1 + \beta_2 post2 + \beta_3 treat1 + \beta_4 treat2 + \beta_5 post1\_treat1 \\ &+ \beta_6 post1\_treat2 + \beta_7 post2\_treat1 + \beta_8 post2\_treat2 + \varepsilon \qquad (4.1) \end{aligned}$$

where $lnexp$ is log of expenditures, $post$ dummies indicate post-decentralization years, and $treat$ dummies stand for our treatment groups (development and non-establishment). All variables used in estimation have been defined in appendix Table C.6.

Table 4.1: Difference-in-Difference Coefficients

	Control	Treatment1	Treatment2
	Establishment exp.	Development exp.	Non-Establishment exp.
Pre-decentralization (Pre)	β_0	$\beta_0 + \beta_3$	$\beta_0 + \beta_4$
Post-decentralization Year 1 ($Post1$)	$\beta_0 + \beta_1$	$\beta_0 + \beta_1 + \beta_3 + \beta_5$	$\beta_0 + \beta_1 + \beta_4 + \beta_6$
Post-decentralization Year 2 ($Post2$)	$\beta_0 + \beta_2$	$\beta_0 + \beta_2 + \beta_3 + \beta_7$	$\beta_0 + \beta_2 + \beta_4 + \beta_8$
Difference (Post1 - Pre)	β_1	$\beta_1 + \beta_5$	$\beta_1 + \beta_6$
Difference (Post2 - Pre)	β_2	$\beta_2 + \beta_7$	$\beta_2 + \beta_8$
Difference-in-Difference1 $(Post1 - Pre)_T$ - $(Post1 - Pre)_C$		β_5	β_6
Difference-in-Difference2 $(Post2 - Pre)_T$ - $(Post2 - Pre)_C$		β_7	β_8

The coefficients of our interest are β_5, β_6, β_7 and β_8: the first two coefficients represent the impact of decentralization on development expenditures and non-establishment expenditures for post-decentralization year 1 respectively, and the last two represent this impact on development and non-establishment expenditures for post-decentralization year 2 respectively. These coefficients can be cast as the difference in expected expenditures of a particular type between the treatment and control groups before and after a treatment, as shown by Wooldridge (2002, p. 130). Wooldridge maintains that although this estimator has been labeled the difference-in-difference (D-D) estimator in program evaluation literature, it has a long history in analysis of variance as the full-interaction model.

Specifically, each of these coefficients represents the amount that the treatment expenditures (development and non-establishment) went up in the first and second year after decentralization *beyond* the amount that the control expenditures (establishment) went up in the same period.

The results presented in Table C.7 constitute our main result. Estimation results from the basic OLS estimation of Equation 4.1 are displayed in column DD-I. In all specifications, we report standard errors that have been clustered at the district level and are robust to heteroskedasticity.

Next we control for district-specific trends and and sector fixed effects. We utilize the following empirical specification:

$$\begin{aligned} lnexp &= \beta_0 + \beta_1 post1 + \beta_2 post2 + \beta_3 treat1 + \beta_4 treat2 + \beta_5 post1\_treat1 \\ &+ \beta_6 post1\_treat2 + \beta_7 post2\_treat1 + \beta_8 post2\_treat2 \\ &+ \sum_{i=1}^{34} district_i\_post1 + \sum_{i=1}^{34} district_i\_post2 + \sum_{j=1}^{17} sector_j + \varepsilon \end{aligned} \tag{4.2}$$

Column DD-II displays results with controls for district time effects only, while column DD-III includes, in addition to district time effects, sector fixed effects as well. Column DD-IV displays results from our most extensive specification, where we control for district-sector-time effects - in effect, we provide an intercept for each sector in each district in each of the post-decentralization years. Our empirical specification is then:

$$\begin{aligned}
lnexp &= \beta_0 + \beta_1 post1 + \beta_2 post2 + \beta_3 treat1 + \beta_4 treat2 \\
&+ \beta_5 post1_treat1 + \beta_6 post1_treat2 + \beta_7 post2_treat1 \\
&+ \beta_8 post2_treat2 + \sum_{i=1}^{34}\sum_{j=1}^{17} district_i_sector_j_post1 \\
&+ \sum_{i=1}^{34}\sum_{j=1}^{17} district_i_sector_j_post2 + \sum_{j=1}^{17} sector_j + \varepsilon \qquad (4.3)
\end{aligned}$$

4.5.1 Interpretation of Results

The difference-in-difference (D-D) results, presented in Table C.7, are noteworthy for a number of reasons. First, the results are robust to our empirical specifications and hold even for our most extensive specification in DD-IV.

Second, development allocations register a large increase in both post-decentralization years beyond that of the establishment expenditures in the same period - the differential increase is more than two-fold in the first year and is more than five-fold in the second year after decentralization. This increase is statistically significant at the 1% level in all specifications. This suggests that decision-makers in the district have a strong preference for development expenditures, and that this preference is getting reflected in increasing levels of funding for development.

Third, current non-establishment expenses also register a noticeable increase over pre-decentralization levels in both years beyond that of the establishment expenditures over the same period. This differential increase, statistically significant at the 1% level, is, however, much smaller as compared to the one for development allocations - of the order of 10-20% in both years relative to the pre-decentralization baseline year.

The preference for development expenditures may be driven by greater demands from voters for such goods, but may also be explained by political factors. Development expenditures are politically the most visible of government activities and where preferences are most revealed, and these are also the most discretionary (Hussain 2008). It is important to highlight that local government elections were held on non-party basis, and that the nazims were elected through an indirect election. Expenditures on infrastructure represent targeted goods and may represent a credible means of fulfilling preelectoral promises by individual candidates in the absence of political parties which represent the most important vehicles for developing policy reputations (Keefer et al. 2006). Since development expenditures entail the award of contracts for implementation, rent-seeking behaviour could be another reason for why politicians prefer such expenditures.

In ongoing work, we hope to explore the reasons behind the preference for development expenditures. For this we plan to exploit the heterogeneity of treatment effects across different districts that are marked by varying levels of development, demographics, political competition, and existing level of public provision. We can measure the level of public provision both by the coverage of public services and by the stock of public infrastructure.

The increases in both development and non-establishment expenditures partly reflect a secular growth in resource base of the districts[13]. This secular trend is also captured by the year dummies: $post1$ is positive and strongly significant in all empirical specifications, whereas $post2$ is also positive and statistically significant in two of the empirical specifications. This increase in resources reflects increased revenues for the federal and provincial governments in these years by 7.8% in the first year and by 19.4% in the second year in column 1, that are ultimately passed on to the districts through the mechanism

[13]Having worked on provincial and district budgets during 2001-02 at the Punjab Finance Department, the pattern of results presented here rings true to me.

of the Provincial Finance Commission. However, as mentioned earlier, these resources are transferred to the districts as general grants through single-line-item transfers, which implies that districts have complete discretion in utilizing these resources as far as allocations for development and non-establishment are concerned. Thus districts could choose allocation mixes not only between development and non-establishment expenses, but also over individual sectors. The fact that districts took advantage of enhanced discretionary resources by making greater allocations for development expenditures therefore reveals the preferences of political decision-makers at the district level.

4.5.2 Treatment Effect Heterogeneity

We examine variation in each sector separately to explore if our treatments affect various public goods differently. Tables C.8 and C.9 display results from difference-in-difference estimation corresponding to empirical specification DD-I in Table C.7, separately for individual sectors. We present results from 12 of the 17 decentralized sectors as these constitute major sectors. Tables C.10 and C.11 display D-D estimation results for the same sectors, while controlling for district-specific fixed effects for each post-decentralization year. These results correspond to empirical specification DD-II in Table C.7, separately for individual sectors.

The most important thing to notice from these results is that we get considerable heterogeneity in treatment effects across various sectors. Thus some sectors receive much bigger increases in allocations for development expenditures. The biggest gainers for development allocations in the first year after decentralization are: Housing (nine-fold), Administration, Livestock and Electrification (four-fold). For the second year after decentralization, the biggest gainers in terms of development allocations are: Fisheries (twelve-fold), Administration and Livestock (eleven-fold), Housing and Agriculture (eight-fold), while allocations for Industry sector actually go down two-fold.

In terms of allocations for non-establishment operational expenses, it is significant to note that there is no statistically significant change in some sectors. But Livestock and Fisheries sector register a modest differential increase in non-establishment expenditures for both years after decentralization. Non-establishment allocations for housing and administration actually go down in the first year after decentralization, where there is no statistically significant change in non-establishment allocations for these sectors in the second year. In ongoing work, we are examining the sources of this treatment heterogeneity by looking at political and other factors.

We realize that we get massive effects for certain sectors and though we have confidence in the validity of these results, we are further exploring what is driving the results in each sector. Several plausible explanations can be given at this stage to interpret the results. We have mentioned earlier that there is much less differential increase in non-establishment expenditures compared to development expenditure which, by their very nature, constitute lumpy but discretionary investments on new infrastructure projects. Thus, conditional on approval of a development project for a specific sector, we would expect development expenditures for that sector to vary a lot over the years. By contrast, non-establishment expenditures are less discretionary and their allocations have to meet a minimum threshold in order to maintain public services. For instance, in the Fisheries sector development expenditures increase from a negligible amount in the base year to 1.5 and 10.10 million Rupees in the first and second year after decentralization. These are small increases covering only a few projects, but these translate into a D-D coefficient for differential increase in development expenditures of 2.056 in the first year and 11.87 in the second year after decentralization, respectively. A similar story holds for Livestock and a few other sectors. The massive increases in development expenditures for the Administration sector perhaps reflect the need for new administrative infrastructure after decentralization in the form of offices and buildings for the decentralized departments. The construction of nice office buildings may also reflect the political need of elected politicians to demonstrate their new, dominant role with respect to the district bureaucracy. Thus these big coefficients are perhaps not very surprising, but these certainly put the question regarding the impact of decentralization at a deeper level - why do politicians prefer certain sectors over others in allocating development expenditures? This is a question we wish to explore further in ongoing work.

4.5.3 Difference-in-Difference (D-D) for Non-Decentralized Sectors

For comparison, we also conduct difference-in-difference estimation separately for some non-decentralized sectors. Even though such estimation does not have the same causal interpretation as the one for decentralized sectors, we conduct this exercise in order to compare heterogeneity of treatment effects for decentralized sectors with the ones for non-decentralized sectors. We define non-decentralized or provincial sectors as ones which have not been decentralized to the local governments, and where the provincial government has the authority to make allocative decisions.

Table C.12 displays these results for individual non-decentralized, or provincial, sectors. The results suggest that there is substantial variation in expenditure patterns across non-decentralized sectors. The biggest gainer from the non-decentralized sectors, in terms of enhanced expenditures, is the Police sector whose development expenditures increase six-fold and fourteen-fold in the first and second year after decentralization beyond the increase in its establishment expenditures in the same time periods. The non-establishment expenditures for Police sector also register statistically significant differential increases in both years after decentralization.

Health and Livestock sectors are the biggest losers in terms of resource allocation at the provincial level. Whereas there is no statistically significant differential change for non-establishment or development expenditures in Health sector for both years, the Livestock sector at the provincial level displays no statistically significant differential change in development expenditures but displays modest differential increases in non-establishment expenditures for both years. However, the results for Police sector may have been driven by parallel reforms in that sector during this period that witnessed large infusion of funds into the sector under the "Access to Justice Program" funded by the Asian Development Bank. We are exploring this variation in provincial sectors further in ongoing work.

4.6 Difference-in-Difference-in-Difference (D-D-D) Estimation

We further improve our identification of the impact of decentralization by utilizing difference-in-difference-in-difference (D-D-D) estimation. Here we exploit the fact that some sectors have not been completely decentralized and the province has retained certain functions from these sectors. For instance, in the Health sector, whereas districts are responsible for delivering most health services, the province retains some services that include teaching hospitals, professional colleges and institutes, mental hospital, chemical examiner, etc.

For D-D-D estimation, we disaggregate provincial allocations for four sectors over districts and types of expenditures. The sectors include three that have been partly decentralized - Health, Agriculture, and Livestock sectors - and a fourth sector that has not been decentralized - Police. We exploit the variation within the same sector across functions that have been decentralized and those that have not been, by comparing activities that have been decentralized and those that have not been. This strategy ensures that we control for trends and fixed effects that are particular to a specific sector.

Our empirical specification for D-D-D estimation is:

$$\begin{aligned} lnexp &= \beta_0 + \beta_1 post1 + \beta_2 post2 + \beta_3 treat1 + \beta_4 treat2 + \beta_5 dev + \beta_6 post1\_dev \\ &+ \beta_7 post2\_dev + \beta_8 post1\_treat1 + \beta_9 post1\_treat2 + \beta_{10} post2\_treat1 \\ &+ \beta_{11} post2\_treat2 + \beta_{12} treat1\_dev + \beta_{13} treat2\_dev + \beta_{14} post1\_treat1\_dev \\ &+ \beta_{15} post1\_treat2\_dev + \beta_{16} post2\_treat1\_dev + \beta_{17} post2\_treat2\_dev + \epsilon \end{aligned} \tag{4.4}$$

Here dev is a dummy variable representing a decentralized function.

Our difference-in-difference-in-difference (D-D-D) strategy is explained in the Table 4.2 below, which shows how we derive our outcomes of interest from Equation 4.4. As displayed in this table, our primary outcomes of interest are:

1. $post1\_treat1\_dev$ (β_{14}): difference between the amount that development expenditure went up beyond that of the establishment expenditures for decentralized functions in a sector and the amount that development expenditures went up beyond that of the establishment expenditures for non-decentralized functions in the same sector, in the first year after decentralization,

2. $post2\_treat1\_dev$ (β_{16}): difference between the amount that development expenditure went up beyond that of the establishment expenditures for decentralized functions in a sector and the amount that development expenditures went up beyond that of the establishment expenditures for non-decentralized functions in the same sector, in the second year after decentralization,

3. $post1\_treat2\_dev$ (β_{15}): difference between the amount that non-establishment expenditure went up beyond that of the establishment expenditures for decentralized functions in a sector and the amount that non-establishment expenditures went up beyond that of the establishment expenditures for non-decentralized functions in the same sector, in the first year after decentralization,

4. $post2\_treat2\_dev$ (β_{17}): difference between the amount that non-establishment expenditure went up beyond that of the establishment expenditures for decentralized functions in a sector and the amount that non-establishment expenditures went up beyond that of the establishment expenditures for non-decentralized functions in the same sector, in the second year after decentralization.

Table C.13 displays results from D-D-D estimation. Similar to D-D estimation, we apply more and more stringent controls. Thus, in a parallel strategy to the empirical specifications in Table C.7, column DDD-I presents results from our basic specification, column DDD-II controls for district effects, column DDD-III controls for district and sector fixed effects, while column DDD-IV controls for district-sector-time fixed effects. The first four coefficients in Table C.13 are the primary outcomes of interest to us.

The results suggest that there are statistically significant changes for the main outcomes of interest, and the primary finding is that we still get a significant decentralization effect

Table 4.2: Difference-in-Difference-in-Difference Coefficients

	Control		Treatment1		Treatment2	
	Establish		Develop		Non-Estab	
	dec	non	dec	non	dec	non
Pre-dec (Pre)	$\beta_0 + \beta_5$	β_0	$\beta_0 + \beta_3 +$ $\beta_5 + \beta_{12}$	$\beta_0 + \beta_3$	$\beta_0 + \beta_4 +$ $\beta_5 + \beta_{13}$	$\beta_0 + \beta_4$
Post 1 $(Post1)$	$\beta_0 + \beta_1$ $+\beta_5 + \beta_6$	$\beta_0 + \beta_1$	$\beta_0 + \beta_1 + \beta_3$ $+\beta_5 + \beta_6 + \beta_8$ $+\beta_{12} + \beta_{14}$	$\beta_0 + \beta_1$ $+\beta_3 + \beta_8$	$\beta_0 + \beta_1 + \beta_4$ $+\beta_5 + \beta_6 + \beta_9$ $+\beta_{13} + \beta_{15}$	$\beta_0 + \beta_1$ $+\beta_4 + \beta_9$
Post 2 $(Post2)$	$\beta_0 + \beta_2$ $+\beta_5 + \beta_7$	$\beta_0 + \beta_2$	$\beta_0 + \beta_2 + \beta_3$ $+\beta_5 + \beta_7 + \beta_{10}$ $+\beta_{12} + \beta_{16}$	$\beta_0 + \beta_2$ $+\beta_3 + \beta_{10}$	$\beta_0 + \beta_2 + \beta_4$ $+\beta_5 + \beta_7 + \beta_{11}$ $+\beta_{13} + \beta_{17}$	$\beta_0 + \beta_2$ $\beta_4 + \beta_{11}$
$D1$ $(Post1 - Pre)$	$\beta_1 + \beta_6$	β_1	$\beta_1 + \beta_6 +$ $\beta_8 + \beta_{14}$	$\beta_1 + \beta_8$	$\beta_1 + \beta_6 +$ $\beta_9 + \beta_{15}$	$\beta_1 + \beta_9$
$DD1$ ($D1_{dec}$-$D1_{non}$)	β_6		$\beta_6 + \beta_{14}$		$\beta_6 + \beta_{15}$	
$DDD1$ ($DD1_T$-$DD1_C$)			β_{14}		β_{15}	
$D2$ $(Post2 - Pre)$	$\beta_2 + \beta_7$	β_2	$\beta_2 + \beta_7 +$ $\beta_{10} + \beta_{16}$	$\beta_2 + \beta_{10}$	$\beta_2 + \beta_7 +$ $\beta_{11} + \beta_{17}$	$\beta_2 + \beta_7$
$DD2$ ($D2_{dec}$-$D2_{non}$)	β_7		$\beta_7 + \beta_{16}$		$\beta_7 + \beta_{17}$	
$DDD2$ ($DD2_T$-$DD2_C$)			β_{16}		β_{17}	

dev: decentralized sectors (dev = 1); non: non-decentralized sectors (dev = 0). T: Treatment; C: Control

for development expenditures for decentralized functions in the second year after decentralization. The coefficient on $post1\_treat1\_dev$ is modestly positive in three empirical specifications, although not statistically significant at the conventional levels. For the second year after decentralization, development allocations for the decentralized sectors register a big differential increase, from 1.4 times in one specification to 1.7 times in three specifications. All of these increases are statistically significant at the 1% level. In comparison, allocations for non-establishment expenditures for the decentralized sectors register a differential decrease for both post-decentralization years under all empirical specifications, although the coefficients are not statistically significant at the conventional levels.

The results thus provide further evidence in support of the conclusions from the D-D estimation that district governments have tended to favour development expenditures over non-establishment expenditures.

4.6.1 Comparison of D-D for Decentralized and Non-Decentralized Sectors

Table C.14 compares D-D estimation results for decentralized and non-decentralized sectors. This is done separately for sectors in which a subset of the total functions performed in that sector have been decentralized, but where the remaining functions in that sector continue to be non-decentralized, or provincially managed. These sectors include: Agriculture, Health and Livestock. For each of these sectors, the table compares D-D estimation results, corresponding to empirical specification DD-II in Table C.7, conducted separately for functions that have been decentralized and for functions that have not been decentralized. Each observation in this estimation represents expenditure allocation for a specific function within a sector pertaining to a particular district and a particular year[14].

This comparison is worth-doing because it helps us uncover the different priorities of district and provincial governments pertaining to the same sector. The comparison, thus, demonstrate that for all these sectors, district governments make bigger increases in allocations for development expenditures beyond those for establishment expenditures in the same sectors than the province in post-decentralization years. The contrast between district and provincial allocations for development is especially stark in case of Livestock sector in both post-decentralization years, and in the Agriculture sector for the second year after decentralization. For non-establishment expenses, the patterns for allocations between the

[14]The expenditures for the non-decentralized (i.e., the provincial) sectors had been disaggregated by districts prior to this estimation.

province and the districts are broadly similar, although districts seem to be allocating less differential increases for non-establishment expenses in the Health sector than the province.

4.7 Identification of Political Channels

We next explore the political channels of local decision-making to better interpret the considerable heterogeneity of results across sectors, especially since we get the somewhat surprising effect of less increase in social sectors.

Using data from elections, we estimate the probability of reelection of local politicians as a function of provision of different types of goods. We broadly classify the district expenditure allocations into three types comprising administrative, physical infrastructure and social sectors. Physical infrastructure type includes building of roads, electrification, etc.; social sectors include education, health, etc.; while administrative sectors include general administration, revenue and tax administration. These variables are defined as the percentage increase in spending on these sectors over the base year.

As a first stage, we explore the correlation between electoral prospects of the district nazim and the spending on local public goods. We have data on whether the district nazim who won the election during the 2001 elections was reelected during the 2005 elections or not. For each of the post-decentralization years 1 and 2, we determine the percentage increase in district expenditures, aggregated over non-development and development, for broad types of sectors as discussed earlier. We estimate a probit model of the equation:

$$reelection_{it} = \Phi(\alpha_0 + \alpha_1 Physical_{it} + \alpha_2 Social_{it} + \alpha_3 Post1 + \alpha_4 \textbf{Controls}_{it}) \quad (4.5)$$

where $\Phi(\cdot)$ is the cumulative distribution function for the standard normal variable. Our controls include the following variables:

- population density,
- household size,
- literacy ratio,
- proportion of *pacca* housing[15],
- ethnic diversity of the district (*fractionalization*),
- proportion of district population living in urban areas.

We measure ethnic diversity, defined by the $diversity$ variable, in terms of fractionalization index which in return is defined as:

$$fractionalization = 1 - \sum_i (\text{fraction of people speaking a particular language}) \quad (4.6)$$

It measures the probability that any person drawn randomly from the population will not speak the ethnic languages listed in the summation which comprises different self-identified ethnic/linguistic groups (Urdu-speaking, Punjabi-speaking, Pashto-speaking, etc.). It ranges from 0 (complete homogeneity) to 1 (complete heterogeneity). A complete list of variables is provided in Table C.6.

The results of our estimation of Equation 4.5 are shown in Table C.15. The results, robust to different empirical specifications, demonstrate the importance of the provision of physical infrastructure to the reelection prospects of the district nazim *ex post*. Thus, nazims who increased the proportion of investment in physical infrastructure had a higher

[15] It stands for cemented, as opposed to mud and stone, construction and is a proxy for wealth.

probability of being re-elected. Spending increases on physical sectors has thus a significant effect, both economically and statistically (at the 5% level), on the reelection prospects of the nazim, and this effect is robust to the addition of more controls. The coefficient on social sectors is actually negative and is statistically significant at the 10% level in one specification, and at the 5% level in two specifications. This suggests that increases in spending on social sectors do not help the reelection prospects of the district nazim.

One possible explanation of these results is that once the province chooses to dominate provision of social services, it is a rational behaviour on the part of nazims to differentiate themselves by choosing physical infrastructure. So nazims are ex ante differentiating themselves in ways that the voters are reacting to favourably ex post. Thus it appears that some goods are either NOT as important to the local politicians' re-election prospects ex post. Thus the argument is that, while social sectors and physical infrastructure are both politically salient, by accident the social sectors have become centrally politically salient, and for this very reason become locally politically non-salient.

Another possible explanation of the findings is a "substitution effect": a decision by voters and local politicians to concentrate on demanding and supplying LOCAL public good mandates that are underfunded by higher governments. We are currently exploring these explanations.

These results suggesting the importance of physical infrastructure are also supported by other evidence. Thus Akramov et al. (IFPRI 2008) find weak preference matching of voters' preferences and policy priorities of candidates for local government positions in Pakistan. They suggest that policy breakdowns and the inability of politicians to make credible pre-electoral policy promises, which were prevalent during the period before decentralization and which produced low allocations for universal goods (Keefer 2002, 2004;

Keefer and Khemani 2003), are also apparent at the local level after decentralization.

Hussain (2008) finds that local government sectoral priorities are heavily tilted towards the provision of physical infrastructure at the expense of education and health. He suggests that this sectoral prioritization is in part a response to the relatively greater citizen demands for physical infrastructure; in part this is a reflection of the local government electoral structure that gives primacy to village and neighborhood-specific issues; and in part it is a reaction to provincial initiatives in education and health that have taken the political space away from local governments in the social sectors, thereby encouraging them to focus more towards physical infrastructure.

In on-going work, we are attempting to disentangle political channels of accountability at the local level by examining variation in achievement of targets by districts on performance grants from higher governments. These grants were funded through the large infusion of funds to Pakistan after September 11, 2001 (9/11). The random shock of 9/11 was contemporaneous to decentralization reforms in Pakistan. This infusion of funds from external donors flowed to the federal and provincial, but not the district, governments. The provincial governments then passed on these funds to district governments under a system of tied or conditional "vertical" grants. These grants did not form a part of the district budget, but represented additional sources of funding for the districts earmarked for specific activities in specific sectors, mostly education and health in the case of Punjab. Future funding levels for these grants were contingent on achievement of targets for previous years.

We are thus exploiting 9/11 as a second 'natural' experiment, orthogonal to the first one, to make the identification because this second 'experiment' provides an exogenous source of variation in the incentive structure of local decision makers. This variation can possibly occur through two means: by changing the *relative political salience*[16] of goods

[16]We measure political salience of different goods in terms of their contribution to the reelection prospects

in different sectors; or from a decision by local decision-makers to concentrate on *local public good mandates* that are underfunded by higher governments.

We are currently exploring these themes using our third treatment (9/11) at the sector level while also using the substantial demographic and political heterogeneity across the districts.

4.8 Conclusion

One of the fundamental concerns of scholars and development practitioners regarding decentralization has been whether local governments are likely to provide more services to the people. The Pakistan experience presents an interesting case to address this question because it underwent a major exercise in electoral, fiscal and administrative decentralization in 2001.

This study examines the expenditure priorities of local governments, and explores the political channels which determine these expenditure priorities for provision of public goods. We use a novel source of (within-sector) variation, and we find a large decentralization effect that is robust to more and more stringent empirical specification. Thus we find that allocations for development expenditures register a large, and statistically significant, increase in both post-decentralization years. In comparison, the allocations for non-establishment expenditures only register a small increase in both post-decentralization years. The preference for development expenditures seems to be driven by political considerations at the local level. Thus the high-visibility of the "brick and mortar" type development schemes may constitute a credible way to provide direct targeted goods to voters.

But our analysis of the heterogeneity of treatment across sectors seems to point out that

of the local elected politicians.

this trend for greater development allocations is not uniform across all sectors and seems to be driven by a subset of sectors that does not include most of the social sectors. Thus, our analysis seems to suggest that this preference for development allocations may have been driven at the expense of broad universal public services (Keefer 2002, 2004; Keefer and Khemani 2003). Furthermore, our finding of considerable heterogeneity of treatment effect suggests that different public goods may be affected differently by decentralization, perhaps based on political considerations.

We improve identification further with more assumptions by comparing treatment heterogeneity between decentralized and non-decentralized sectors. Here we exploit variation within the same sectors across functions that have been decentralized and those that have not been, by comparing activities that have been decentralized and those that have not been. This gives us difference-in-difference-in-difference (D-D-D) estimates of the impact of decentralization, where we still find a substantial, and statistically significant, effect on development expenditures, but not on non-establishment expenditures.

In ongoing work, we are examining variation in budget shares of different sectors pre- and post-decentralization, and we are also extending the analysis to a sub-sector level. We are also exploring channels of local political accountability by distinguishing between goods that have different local political salience; and exploring if local politicians differentiate themselves by allocating more to goods that have greater local political salience in ways that the voters are reacting to favorably ex post. We are also exploring an alternative explanation for this pattern of expenditure allocations in terms of a possible "substitution effect". In future work, we plan to explore impacts of decentralization on socioeconomic outcomes.

The findings presented in this study have important policy implications for optimal design of decentralization reform and for improving local government accountability. One of these points to the need for ensuring coherence in the key accountability relationships. Thus, shared mandates for delivering the same services between different levels of governments seem to generate perverse incentives for politicians and should be avoided. Other findings also suggest that decentralization *per se* may not be enough to bring about a radical improvement in public provision in the hitherto neglected social sectors, and that the key to unraveling the mystery of when decentralization generates a desired impact on public provision may lie in the political market imperfections that govern the system. This is a theme we wish to explore in future work.

Chapter 5

Conclusions and Further Work

We have analyzed a variety of economic questions throughout this thesis covering the labour market, volunteering market, and the political market. This dissertation demonstrates uses of consistent microeconometric estimation strategies for estimating key economic outcomes by overcoming the econometric problems of correlated unobserved effects, endogeneity, misspecification and reflection. We particularly focus attention on the channels that affect economic behaviour in each case, and its links to policy.

We demonstrate the importance of social networks on labour market outcomes in the second chapter. We make a distinction between weak and strong ties and demonstrate that weak ties matter in employment outcomes, but strong ties do not. We also show that the effect of social networks varies along several interesting dimensions that include age, gender, union coverage, high school completion and skill level.

We explore the determinants of volunteering behaviour in the second study (Chapter 3), and specifically examine the effect of employment on levels of formal and informal volunteering. We provide strong evidence that employment status has a substantial effect on formal volunteering behaviour, but a weaker effect on informal volunteering.

The third study in this dissertation (Chapter 4) uses constructed data from Pakistan to estimate the impact of political decentralization. The contribution of the third essay is to demonstrate a large decentralization effect that is robust to the introduction of numerous controls. It also shows an interesting pattern of heterogeneity of treatment effects: the decentralization effect is only driven by a specific subset of sectors that does not include the social sectors. There is some evidence to suggest that this pattern is driven by the rational response of local politicians to the peculiar incentive structures introduced by higher levels of government.

5.1 Policy Implications

This dissertation contributes to improving our knowledge of the channels through which policy interventions can have an impact. The results of Chapter 2 and Chapter 3 suggest that policy interventions that utilize the social and employment networks of individuals can have a significant "social multiplier" effect. Chapter 2 suggests that policy interventions that augment the weak ties of individuals can have substantial effects on their labour market outcomes.

Chapter 3 suggests that the traditional concern that work time squeezes out time spent on volunteering may be misplaced, and in fact the two may be positively related. Thus, working for pay does not necessarily substitute for working without pay, or volunteering. Our results thus suggest that policy interventions aimed at generating employment opportunities can have spillover effects in terms of encouraging volunteering behaviour that, in turn, may generate positive social and economic outcomes. However, our findings also suggest that the relationship between employment and volunteering behaviour may depend on the precise intermediation channels through which working individuals get connected by volunteer organizations, a theme that needs to be further explored.

Chapter 4 in examining the impact of decentralization also explores what public goods are affected by it. Our results throw some light on the political channels through which decentralization has an effect. Our findings on the political channels suggest that decentralization and other governance reforms need to take account of the incentive structures facing local politicians, and the strategic interaction between different levels of government. The results suggest the need for ensuring coherence in key accountability relationships for optimal design of decentralization reform. Thus, shared mandates between governments at different levels for delivering the same function do not seem to provide appropriate incentives for optimal resource allocation. To our knowledge, these results are new to the literature, and are particularly relevant for policy purposes.

5.2 Ongoing and Future Work

There are a number of interesting questions that extend from this dissertation that we intend to address in our ongoing and future research. Thus, because our data come from unique settings, we want to explore if our results are broadly representative. This is especially true for the data used in the first two essays where it comes from a low-income group of people located in the Atlantic region of Canada, and we find it worth examining if the patterns of results we get are representative of other classes and regions.

Currently, we have only utilized data from the first wave of a three-wave survey in Chapter 2. We will later utilize the panel nature of our data and explore the dynamic interaction of social networks and labour market outcomes. We also wish to explore whether social networks have a significant impact on subsequent welfare participation of individuals. We will examine this question by exploring the relationship between changes in social networks and changes in welfare participation.

Regarding volunteering, we are currently exploring in ongoing work the channels through which employment has an effect on volunteering behaviour. Further possible work will involve estimating the equation for time spent on volunteering jointly with one for time spent on market labour. It will also involve exploring a dynamic specification of volunteering behaviour by taking advantage of the panel nature of the data.

For future work, we may combine the analysis in the first two papers by exploring the relationship between social networks and philanthropic behaviour including volunteering. The recent work of Apinunmahakul and Devlin (2008) provides a theoretical framework for this relationship and the authors then conduct an empirical investigation of the link between networks and private philanthropy using the Canadian National Survey of Giving, Volunteering and Participating. Given the greater degree of exogenous variation in our data set, conducting a similar empirical investigation by using our data set seems quite promising.

For our decentralization work, we are currently examining variation in budgetary shares of different sectors due to decentralization. We are also exploring using the semi-parametric technique of quantile regression to characterize the conditional distribution of budgetary expenditures. We are also extending the analysis to a sub-sector level to explore precisely what public goods are affected by decentralization and why. This work will involve exploring channels of local political accountability by distinguishing between goods that have different local political salience, and examining their relationship with electoral mechanisms. Further possible work will involve exploring the impact of decentralization on socioeconomic outcomes such as income levels, and education and health indicators.

References

Ahmad, J., S. Devarajan, S. Khemani, and S. Shah (2004) "Decentralization and Service Delivery," Mimeo, World Bank.

Ajmal, M., and S. Bari (2006) "Understanding the role of political dynasties: Local government elections 2005," Working paper, Pattan Development Organization, Islamabad, Pakistan.

Akramov, K. T., S. Qureshi, R. Birner, and B. H. Khan (2008) "Decentralization, Local Government Elections and Voter Turnout in Pakistan," IFPRI Discussion Paper 00754.

Alesina, A., R. Baqir, and W. Easterly (1999) "Public Goods and Ethnic Divisions," *Quarterly Journals of Economics* 114(4).

Altonji, J., and L. Segal (1996) "Small-sample Bias in GMM Estimation of Covariance Structures," *Journal of Business and Economic Statistics*, 14, 353-366.

Anderson, T.W. and H. Rubin (1949) "Estimation of the parameters of a single equation in a complete system of stochastic equations," *Annals of Mathematical Statistics* 20: 46-63.

Andreoni, J. (1998) "Toward a theory of charitable fundraising," *Journal of Political Economy* 106, 1186-1213.

Angrist, J. D. and G. Imbens (1994) "Identification and Estimation of Local Average Treatment Effects," *Econometrica* 62(2): 467-475.

Angrist, J. D. and G. Imbens (1995) "Two Stage Least Squares Estimation of Average Causal Effects in Models with Variable Treatment Intensity," *Journal of the American Statistical Association* 90(430): 431-442.

Antoni, G. D. (2009) "Intrinsic vs. Extrinsic Motivations to Volunteer and Social Capital Formation," *Kyklos* 62(3): 359-370.

Apinunmahakul, A., and R. A. Devlin (2008) "Social networks and private philanthropy," *Journal of Public Economics* 92(1-2), 309-328.

Apinunmahakul, A., V. Barham, and R. A. Devlin (2009) "Charitable giving, volunteering and the paid labor market," *Nonprofit and Voluntary Sector Quarterly* 38(1), 77-94.

Applied Economic Research Centre (AERC) (1990) "Local Government Administration in Pakistan," Karachi: unpublished manuscript.

Arrow, K. J. (1974) "Gifts and Exchanges," *Philosophy and Public Affairs*, 1, 343-62.

Azfar, O., S. Kahkonen, and P. Meagher (2001) "Conditions for effective decentralized governance: A synthesis of research findings," IRIS Working Paper No. 256, University of Maryland.

Banerjee, A., and R. Somanathan (2006) "The Political Economy of Public Goods: Some Evidence from India," Working Paper, Massachusetts Institute of Technology.

Bardhan, P. and D. Mookherjee (2000) "Capture and Governance at Local and National Levels," *American Economic Review* 90:2, pp. 135-39.

Bardhan, P. and D. Mookherjee (2006) "Decentralisation and Accountability in Infrastructure Delivery in Developing Countries," *The Economic Journal* 116 (January): 101-27.

Barro, R. (1973) "The control of politicians: an economic model," *Public Choice* 14:19-42.

Baum, C. F., M. E. Schaffer, and S. Stillman (2007) "Enhanced routines for instrumental variables/GMM estimation and testing," Boston College Economics Working Paper No. 667.

Baum, C. F., M. E. Schaffer, and S. Stillman (2003) "Instrumental variables and GMM: Estimation and testing," *Stata Journal* 3: 1-31.

Baum, C. F. (2006) "An Introduction to Modern Econometrics Using Stata," College Station, TX: Stata Press.

Bayer, P., S. Ross, and G. Topa (2008) "Place of Work and Place of Residence: Informal Hiring Networks and Labour Market Outcomes," *Journal of Political Economy*, 116:6, 1150-1196.

Beaman, L. A. (2008) "Social Networks and the Dynamics of Labor Market Outcomes: Evidence from Refugees Resettled in the U.S.," Northwestern University. Mimeo.

Becker, G. (1974) "A Theory of Social Interactions," *Journal of Political Economy*, Vol. 82, 1063-1093.

Becketti S., W. Gould, L. Lillard, and F. Welch (1988) "The Panel Study of Income Dynamics after Fourteen Years: An Evaluation," *Journal of Labor Economics* 6: 472-492.

Bertrand M., E. Duflo, and S. Mullainathan (2004) "How Much Should We Trust Difference-in-Differences Estimates?" *Quarterly Journal of Economics*, 119, pp. 249-75.

Bertrand, M., E. F. P. Luttmer, and S. Mullainathan (2000) "Network effects and welfare cultures," *Quarterly Journal of Economics* 115, 1019-1055.

Besley T. and S. Coate (2003) "Centralized versus Decentralized Provision of Local Public Goods: A Political Economy Analysis," *Journal of Public Economics* 87(1).

Besley T. and A. Case (1995) "Incumbent Behavior: Vote Seeking, Tax Setting and Yardstick Competition," *American Economic Review*, 85, 1: 25-45.

Besley T. and S. Coate (1992) "Workfare versus Welfare: Incentive Arguments for Work Requirements in Poverty Alleviation Programs," *American Economic Review* 82, 1, March 1992: 249-261.

Besley, T. and R. Kanbur (1993) "The Principles of Targeting," In M. Lipton and J. Gaag (eds.), Including the Poor. Washington, DC: The World Bank, pp. 67-90.

Birner, R., S. Qureshi, K. Akramov and B. H. Khan (2006) "Voice and votes: Does political decentralization work for the poor and for women?," *Pakistan Development Review* 45 : 4 Part II (Winter 2006) pp. 1345-1360.

Boorman, S. A. (1975) "A combinatorial optimization model for transmission of job information through contact networks," *Bell Journal of Economics* 6, 216-249.

Borjas, G. (1992) "Ethnic Capital and Intergenerational Mobility," *Quarterly Journal of Economics*, 107, 123-150.

Bound, J., D. A. Jaeger, and R. Baker (1995) "Problems with instrumental variables estimation when the correlation between the instruments and the endogeneous explanatory variable is weak," *Journal of the American Statistical Association* 90: 443-450.

Cahuc, P. and F. Fontaine (2009) "On the Efficiency of Job Search with Social Networks," *Journal of Public Economic Theory* 11(3), pp. 411-439.

Calvo-Armengol, A. and M. O. Jackson (2004) "The effects of social networks on employment and inequality," *American Economic Review* 94, 426-454.

Calvo-Armengol, A. and M. O. Jackson (2007) "Networks in labour markets: Wage and employment dynamics and inequality," *Journal of Economic Theory* 132, 27-46.

Card, D. and A. Krueger (1992) "Does school quality matter? Returns to education and the characteristics of public schools in the United States," *Journal of Political Economy* 100(1), 1-40.

Card, D. (1999) "The Causal Effect of Education on Earnings," in O. Ashenfelter and D. Card, eds., Handbook of labor economics, Vol. 3A. Amsterdam: Elsevier Science, 1999, pp. 1801-63.

Case, A. and L. Katz (1991) "The Company You Keep: The Effects of Family and Neighborhood on Disadvantaged Youth," Working Paper No. 3705, National Bureau of Economic Research.

Chaudhuri, S. (2006) "What difference does a constitutional amendment make? The 1994 Panchayati Raj Act and the attempt to revitalize rural local government in India," In Decentralization in Developing Countries: A Comparative Perspective (eds. P. Bardhan and D. Mookherjee) MIT Press.

Cheema, A. and S. Mohmand (2006) “Bringing Electoral Politics to the Doorstep: Who Gains who Loses?,” Mimeo, Lahore University of Management Sciences.

Cheema, A., Khwaja, A.I., and Qadir, A. (2006) “Decentralization in Pakistan: Context, Content and Causes” in D. Mookherjee and P. Bardhan (eds.) Local Governments and Decentralization in Developing Countries, MIT Press.

Chernozhukov, V, and C. Hansen (2008) “The Reduced Form: A Simple Approach to Inference with Weak Instruments,” *Economics Letters*, 100: 68-71.

Cragg, J. G., and S. G. Donald (1993) “Testing identifiability and specification in instrumental variables models,” *Econometric Theory* 9: 222-240.

Crook, R.C. and J. Manor (1998) “Democracy and decentralization in South Asia and West Africa: Participation, accountability and performance,” Cambridge: Cambridge University Press.

Crook, R.C. and A. S. Sverrisson (2001) “Decentralization and poverty alleviation in developing countries: A comparative analysis or Is West Bengal unique?” IDS Working Paper No. 130, Institute of Development Studies, Brighton, UK.

Davidson, R. and M. MacKinnon (1993) “Estimation and Inference in Econometrics,” New York: Oxford University Press.

Day, K.M., and R.A. Devlin (1996) “Volunteerism and Crowding Out: Canadian Econometric Evidence,” *Canadian Journal of Economics* 29(1): 37-53.

Day, K.M., and R.A. Devlin (1998) “The payoff to work without pay: volunteer work as an investment in human capital,” *Canadian Journal of Economics* 31(5): 1179-91.

Deaton, A. (1997) “The Analysis of Household Surveys,” Washington D.C.: The World Bank.

van Dijk, J., and R. Boin (1993) “Volunteer Labor Supply in the Netherlands,” *De Economist*, 141, No. 3, 402-418.

Duflo, E. and E. Saez (2003) “The Role of Information and Social Interactions in Retirement Plan Decisions: Evidence from a Randomized Experiment”, *Quarterly Journal of Economics*, 118(3): 815-842.

Dufour, J.M. (2003) “Identification, Weak Instruments and Statistical Inference in Econometrics,” *Canadian Journal of Economics*, 36(4): 767-808.

Easterly, W. (2003) “The Political Economy of Growth Without Development: A Case Study of Pakistan” in D. Rodrik (eds.) ‘In Search of Prosperity: Analytical Narratives on Economic Growth’. Princeton University Press, New Jersey.

Evans, W. N., Oates, W. E., and Schwab, R. M. (1992) “Measuring peer group effects: A study of teenage behavior,” *Journal of Political Economy*, 100(5), 966-991.

Faguet, J.P. (2004) “Does Decentralization increase Government Responsiveness to Local Needs: Evidence from Bolivia,” *Journal of Public Economics* 88 (3-4), pp. 867-893

Fontaine F. (2008) “Why are similar workers paid differently? The role of social networks,” *Journal of Economic Dynamics and Control* 32(12), 3960-3977.

Freeman, R. (1993) “Me, Give to Charity - Well, If You Ask,” Paper presented at the Conference on Happiness, Fairness, and Markets, London School of Economics.

Freeman R. (1997) “Working for nothing: the supply of volunteer labor,” *Journal of Labor Economics*. 15:140-67.

Fukuyama, F. (1995) “Trust: The Social Virtues and the Creation of Prosperity,” New York, Free Press.

Gaviria, A., Raphael, S. (2001) “School-based peer effects and juvenile behavior,” *Review of Economics and Statistics*, 83(2), pp. 257-268.

Gazdar, H. (2000) “State, Community, And Universal Education: A Political Economy of Public Schooling in Rural Pakistan,” Typescript, Asia Research Centre, London School of Economics.

Gomez, R. and M. Gunderson (2003) “Volunteer Activity and the Demands of Work and Family Life,” *Industrial Relations* 58: 573-591.

Goss, K. (1999) “Volunteering and the Long Civic Generation,” *Nonprofit and the Voluntary Sector Quarterly* 28: 378-415.

Granovetter, M. (1973) “The Strength of Weak Ties,” *American Journal of Sociology*, 78, 1360-1380.

Granovetter, M. (1974) “Getting a job: A study of contacts and careers,” Chicago: University of Chicago Press.

Greene, W.H. (2003) “Econometric Analysis,” 5th ed. Prentice Hall, Upper Saddle River, NJ.

Greenwood, J., Nicholson, C., Gyarmati, D., Kyte, D., MacInnis, M., and Ford, R. (2003) “The community employment innovation project: Design and implementation,” Ottawa: Social Research and Demonstration Corporation.

Guess, G. M. (2005) “Comparative decentralization lessons from Pakistan, Indonesia, and the Philippines,” *Public Administration Review* 65(2): 217-230.

Gyarmati, D., and Kyte, D. (2003) “Social capital, network formation and the Community Employment Innovation Project,” *Horizons*, 6(3)-218.

Gyarmati, D., de Raaf, S., Nicholson, C., Kyte, D., and MacInnis, M. (2006) “Testing a community-based jobs strategy for the unemployed: Early impacts of the Community Employment Innovation Project,” Ottawa: Social Research and Demonstration Corporation.

Gyarmati, D., de Raaf, S., Nicholson, C., Palameta, B., Hui, T., and MacInnis, M. (2007) "Improving skills, networks, and livelihoods through community-based work: Three year impacts of the Community Employment Innovation Project," Ottawa: Social Research and Demonstration Corporation.

Gyarmati, D., de Raaf, S., Nicholson, C., Palameta, B., Hui, T., Kyte D, and MacInnis, M. (2008) "Engaging Communities in Local Development: Measuring the effects of the Community Employment Innovation Project," Ottawa: Social Research and Demonstration Corporation.

Hackl, F., M. Halla and G. Pruckner (2007) "The Fallacy of the Good Samaritan: Volunteering as a Weird Way of Making Money," *Kyklos* 60(1): 77-104.

Hall, P., and J. Horowitz (1996) "Bootstrap Critical Values for Tests Based on Generalized-method-of-moments Estimators," *Econometrica*, 64, 891-916.

Hall, M., T. Knighton, P. Reed, P. Bussiere, D. Macrae, and P. Bowen (1998) "Caring Canadians, Involved Canadians: Highlights from the 1997 National Survey of Giving, Volunteering and Participating," Ottawa: Statistics Canada.

Hall, M., L. McKeown, and K. Roberts (2001) "Caring Canadians, Involved Canadians: Highlights from the 2000 National Survey of Giving, Volunteering and Participating," Ottawa: Statistics Canada.

Hall, M., D. Lasby, S. Ayer, and W. D. Gibbons (2009) "Caring Canadians, Involved Canadians: Highlights from the 2007 Canada Survey of Giving, Volunteering and Participating," Ottawa: Statistics Canada.

Hansen, L. (1982) "Large Sample Properties of Generalized Method of Moments Estimators," *Econometrica*, 50, 1029-1054.

Hausman, J.A. (1978) "Specification tests in econometrics," *Econometrica* 46, 1251-1271.

Hayashi, F. (2000) "Econometrics," Princeton, NJ: Princeton University Press.

Heckman, J.J. (2001) "Accounting for Heterogeneity, Diversity and General Equilibrium in Evaluating Social Programs," Economic Journal, 111, 654-699.

Hindricks, J. and B. Lockwood (2005) "Decentralization and electoral accountability: Incentives, separation, and voter welfare," Warwick Economics Research Paper No. 729, Warwick University.

Holzer, H. (1988) "Search Methods Used by Unemployed Youth," *Journal of Labour Economics* 6(1), 1-20.

Hussain, Z. (2008) "Devolution, Accountability and Service Delivery: Some Insights from Pakistan," World Bank Mimeo.

Hussain, I. (1999) "Pakistan: The Economy of an Elitist State," Karachi: Oxford University Press.

Imbens, G. and J. Wooldridge (2009) “A Course in Panel Data Econometrics,” Toronto: Canadian Economics Association and Canadian Labour Market and Skills Research Network.

Independent Sector (1986) “Americans Volunteer 1985,” Washington, D.C.: Independent Sector.

Ioannides, Y. M. and L. D. Loury (2004) “Job information networks, neighborhood effects, and inequality,” *Journal of Economic Literature* 42, 1056-1093.

Isham, J., D. Narayan, and L. Pritchett (1995) “Does Participation Improve Performance? Establishing Causality with Subjective Data,” *The World Bank Economic Review*, Vol. 9, No. 2, 175-200.

Jackson, M. O. (2009) “Networks and economic behavior,” *The Annual Review of Economics*, Vol. 1, 489-513.

Johnson, C. (2003) “A model of social capital formation,” Ottawa: Social Research and Demonstration Corporation.

Johnson, M., K. Foley, and G. Elder (2004) “Women’s Community Service, 1940-1960: Insights from a Cohort of Gifted American Women,” *Sociological Quarterly* 45: 45-66.

Katz, L. F., Kling, J. R., and Liebman, J. B. (2001) “Moving to opportunity in Boston: Early results of a randomized mobility experiment,” *Quarterly Journal of Economics*, 116(2), 607-654.

Keefer, P., A. Narayan and T. Vishwanath (2006) “The Political Economy of Decentralization in Pakistan,” in D. Mookherjee and P. Bardhan (eds.) ‘Decentralization to Local Governments in Developing Countries: A Comparative Perspective’, MIT Press.

Keefer, P. (2002) “Clientelism, Credibility and Democracy,” World Bank, Development Research Group, Washington, D.C.

Keefer, P. (2004) “What does Political Economy tell us about Economic Development - and Vice Versa?” World Bank Policy Research Working Paper No. 3250. Washington D.C.

Keefer, P. and S. Khemani (2003) “Democracy, Public Expenditures, and the Poor,” World Bank Policy Research Working Paper 3164.

Kleibergen, F. (2007) “Generalizing weak instrument robust IV statistics towards multiple parameters, unrestricted covariance matrices and identification statistics,” *Journal of Econometric*s 139, 181-216.

Kleibergen, F., and R. Paap (2006) “Generalized Reduced Rank Tests Using the Singular Value Decomposition,” *Journal of Econometrics*, 133, 97-126.

Kohut, Andrew (1997) “Trust and Citizen Engagement in Metropolitan Philadelphia: A Case Study,” Philadelphia: Pew Research Centre for the People and the Press.

Lasby, D. (2004) “The Volunteer Spirit in Canada: Motivations and Barriers,” Toronto: Canadian Center for Philanthropy.

Laschever, R. (2005) “The doughboys network: Social interactions and labor market outcomes of World War I veterans,” University of Illinois Working Paper, Champaign-Urbana.

Lin, N. (1982) “Social Resources and Instrumental Action,” p. 131-45 in *Social Structures and Network Analysis*, edited by P.V. Marsden and N. Lin. Beverly Hills, CA: Sage Publications.

Litvack, J. and J. Seddon (Eds.) (1999) “Decentralization Briefing Notes,” Washington DC: World Bank Institute.

Ludwig, J., G.J. Duncan, and P. Hirschfield (2001) “Urban Poverty and Juvenile Crime: Evidence from a Randomized Housing-Mobility Experiment,” *Quarterly Journal of Economics*, CVI (2001), 655-681.

Manning, N., D. Porter, J. Charlton, M. Cyan and Z. Hasnain (2003) “Devolution in Pakistan - Preparing for Service Delivery Improvements”, Pakistan: World Bank.

Manski, C. (1993) “Identification of Endogenous Social Effects: The Reflection Problem,” *Review of Economic Studies*, LX, 531-542.

Manski, C. (2000) “Economic Analysis of Social Interactions,” *Journal of Economic Perspectives*, XIV, 115-137.

Menchik, Paul and Burton Weisbrod (1987) “Volunteer Labor Supply,” *Journal of Public Economics* 32: 159-183.

Miguel, E. and M. Kremer (2002) “Why Don’t People Take Their Medicine? Experimental Evidence from Kenya,” mimeo, University of California, Berkeley and Harvard University.

Montgomery, J. D. (1991) “Social networks and labour market outcomes: Toward an economic analysis,” *American Economic Review* 81, 1408-1418.

Mortensen, D. and T. Vishwanath (1994) “Personal contacts and earnings. It is who you know!,” *Labour Economics* 1, 187-201.

Munshi, K. (2003) “Networks in the modern economy: Mexican migrants in the US labour market,” *Quarterly Journal of Economics* 118, 549-599.

Musick, M., and J. Wilson (2008) “Volunteers: A Social Profile,” Bloomington, IN, Indiana University Press.

Myerson, R. (2006) “Federalism and incentives for success of democracy,” *Quarterly Journal of Political Science* 1:3-23.

Myerson, R. (2009) “Constitutional structures for a strong democracy: Considerations on the Government of Pakistan,” Mimeo, University of Chicago.

Newey, W.K. and D. McFadden (1994) “Large sample estimation and hypothesis testing,” in: Engle, R.F., McFadden, D.L. (Eds.), Handbook of Econometrics. Elsevier Science B.V., Amsterdam, pp. 2111-2245, Chapter 36.

Oates, W.E. (1999) “An Essay on Fiscal Federalism,” *Journal of Economic Literature*, Vol. 37, No.3, pp. 1120-1149.

Oesterle, S., M. K. Johnson, and J. Mortimer (2004) “Volunteering during the Transition to Adulthood: A Life-Course Perspective,” *Social Forces* 82: 1123-1151.

Ostrom, E. (1990) “Governing the Commons: The Evolution of Institutions for Collective Action,” Cambridge, Cambridge University Press.

Pande, R., and C. Udry (2005) “Institutions and Development: a View From Below,” Mimeo Yale.

Persson, T. and G. Tabellini (2000) “Political Economics: Explaining Public Policy”, Cambridge, MA: The MIT Press.

Policy Research Initiative (2003) “Social capital: Building on a network-based approach,” Ottawa: Policy Research Initiative.

Profile of Illinois: An Engaged State. (2001) Report of the Illinois Civic Engagement Project http://civic.uis.edu/report/release.html.

Prouteau, L. and F.-C. Wolff (2006) “Does voluntary work pay off in the labor market?” *Journal of Socio-Economics* 35: 992-1013.

Prouteau, L. and F.-C. Wolff (2008) “On the relational motive for volunteer work,” *Journal of Economic Psychology* 29(3), 314-335.

Przeworski, A., S. Stokes and B. Manin (eds) (1999) “Democracy, Accountability and Representation,” Cambridge University Press.

Putnam, R.D., (1993) “Making Democracy Work: Civic Traditions in Modern Italy,” Princeton University Press, Princeton.

Ravanera, Z., R. Beaujot, and F. Rajulton (2002) “The Family and Political Dimensions of Social Cohesion: Analyzing the Link Using the 2000 National Survey on Giving, Volunteering and Participating,” Paper presented at the Annual Meeting of the Canadian Population Society, Toronto. [http://www.ssc.uwo.ca/sociology/popstudies/dp/do02-07.pdf].

Robinson, John and Geoffrey Godbey (1997) “Time for Life: The Surprising Ways Americans Use Their Time,” University Park: Pennsylvania State University Press.

Rockefeller Brothers Fund (1986) “Report on Charitable Giving,” New York: Rockefeller Brothers Fund.

Rose-Ackerman, S. (1982) “Charitable Giving and Excessive Fundraising,” *Quarterly Journal of Economics*, 97, 195-212.

Sacerdote, B. (2001) “Peer effects with random assignment: Results for Dartmouth roommates,” *Quarterly Journal of Economics*, 116(2), 681-704.

Seabright, P. (1996) “Accountability and decentralization in government: An incomplete contracts model,” *European Economic Review* 40(1): 61-89.

Shah, A. (ed.) (2006) “Local governance in developing countries,” Public Sector Governance and Accountability Series. Washington, DC: World Bank.

Shea, J. (1997) “Instrument relevance in multivariate linear models: a simple measure,” *Review of Economics and Statistics* 49(2): 348-352.

Smith, J. D. (1998) “The 1997 National Survey of Volunteering,” London: National Centre for Volunteering.

Smock, K. (2004) “Democracy in Action: Community Organizing and Urban Change,” New York: Columbia University Press.

Stata (2005) “Negative and missing R-squared for 2SLS/IV, ”Retrieved on August 10, 2009. www.stata.com/support/faqs/stat/2sls.html.

Statistics Canada (2006) “2001 community profiles,” Retrieved from www12.statcan.ca/english/profil01/CP01/index.cfm?lang=E.

Stigler, G. (1957) “The tenable range of functions of local government. In Federal expenditure policy for economic growth and stability,” ed. Joint Economic Committee, Subcommittee on Fiscal Policy, U.S. Congress, 213-219. Washington, DC: U.S. Government Printing Office.

Stock, J. H., and J. H. Wright (2000) “GMM with Weak Identification,” *Econometrica* 68(5): 1055-1096.

Stock, J. H. and M. Yogo (2005) “Testing for weak instruments in linear IV regression,” in ‘Identification and Inference for Econometric Models: Essays in Honor of Thomas Rothenberg’, ed. D. W. K. Andrews and J. H. Stock, 80-108. Cambridge: Cambridge University Press.

Stokes, S.C. (2007) “Models of programmatic and clientelistic parties,” In Oxford handbook of comparative politics, ed. Carles Boix and Susan Stokes. New York: Oxford University Press.

Sugden, R. (1984) “Reciprocity: The Supply of Public Goods through Voluntary Contributions,” *Economic Journal*, 94, 772-87.

Sundeen, R. and S. Raskoff (1994) “Volunteering among Teenagers in the United States,” *Nonprofit and Voluntary Sector Quarterly* 23: 383-403.

Taniguchi, H. (2006) “Men’s and Women’s Volunteering: Gender Differences in the Effects of Employment and Family Characteristics,” *Nonprofit and Voluntary Sector Quarterly* 35(1): 83-101.

Tassier, T. (2006) “Labor Market Implications of Weak Ties,” *Southern Economic Journal* 72(3), 704-719.

Tendler, J. (1997) “Good government in the tropics. John Hopkins Studies in Development,” Baltimore and London: John Hopkins University Press.

Terza, J.V., A. Basu, P. Rathouz (2008) “Two-stage residual inclusion estimation: Addressing endogeneity in health econometric modeling,” *Journal of Health Economics* 27: 531-543.

Thomas, A. and H. Finch (1990) “On Volunteering: A Qualitative Research Study,” Unpublished manuscript. London: London School of Economics.

Topa, G. (2001) “Social interactions, local spillovers and unemployment,” *Review of Economics Studies* 68, 261-295.

Triesman, D. (2002) “Decentralization and the Quality of Government,” Mimeo, University of California, Los Angeles.

US Bureau of Labor Statistics (2005) “Volunteering in the United States,” Washington, D.C.: Government Printing Office.

Vaillancourt, F. (1994) “To volunteer or not: Canada, 1987,” *Canadian Journal of Economics* 27, 813-26.

Vaillancourt, F., and M. Payette (1986) “The supply of volunteer work: the case of Canada,” *Journal of Voluntary Action Research* 15(4), 45-56.

Wabha, J. and Y. Zenou (2005) “Density, social networks and job search methods: Theory and application to Egypt,” *Journal of Development Economics* 88, 443-473.

Webb, N., R. Abzug (2008) “Do Occupational Group Members Vary in Volunteering Activity?” *Nonprofit and Voluntary Sector Quarterly* 37(4): 689-708.

Weisbrod, B. A. (1975) “Towards a Theory of the Voluntary Non-Profit Sector in a Three-Sector Economy,” in Phelps, E. S. (ed.) ‘Altruism, Morality, and Economic Theory’, New York, Russell Sage.

White, H. (1994) “Estimation, Inference and Specification Analysis,” Cambridge University Press, Cambridge.

Wilson, J., and Musick, M. (1997) “Who cares? Toward an integrated theory of volunteer work,” *American Sociological Review*, 62: 694-713.

Woolcock, M. (2001) “The place of social capital in understanding social and economic outcomes,” *Isuma*, 2 (1).

Wooldridge, J.M. (2002) “Econometric Analysis of Cross Section and Panel Data,” MIT Press, Cambridge, MA.

World Bank (2003) “Making Services Work for Poor People: World Development Report 2004,” Washington D.C.

World Bank (2001) “Indonesia Poverty Report,” The World Bank: Washington D.C

Zaidi, S. A. (2005) “The political economy of decentralization in Pakistan,” Transversal Theme on Decentralization and Social Movement, Working Paper No. 1, Sustainable Development Policy Institute, Islamabad, and Development Study Group, Zurich.

Appendix A

Appendix of Tables and Figures for Chapter 2

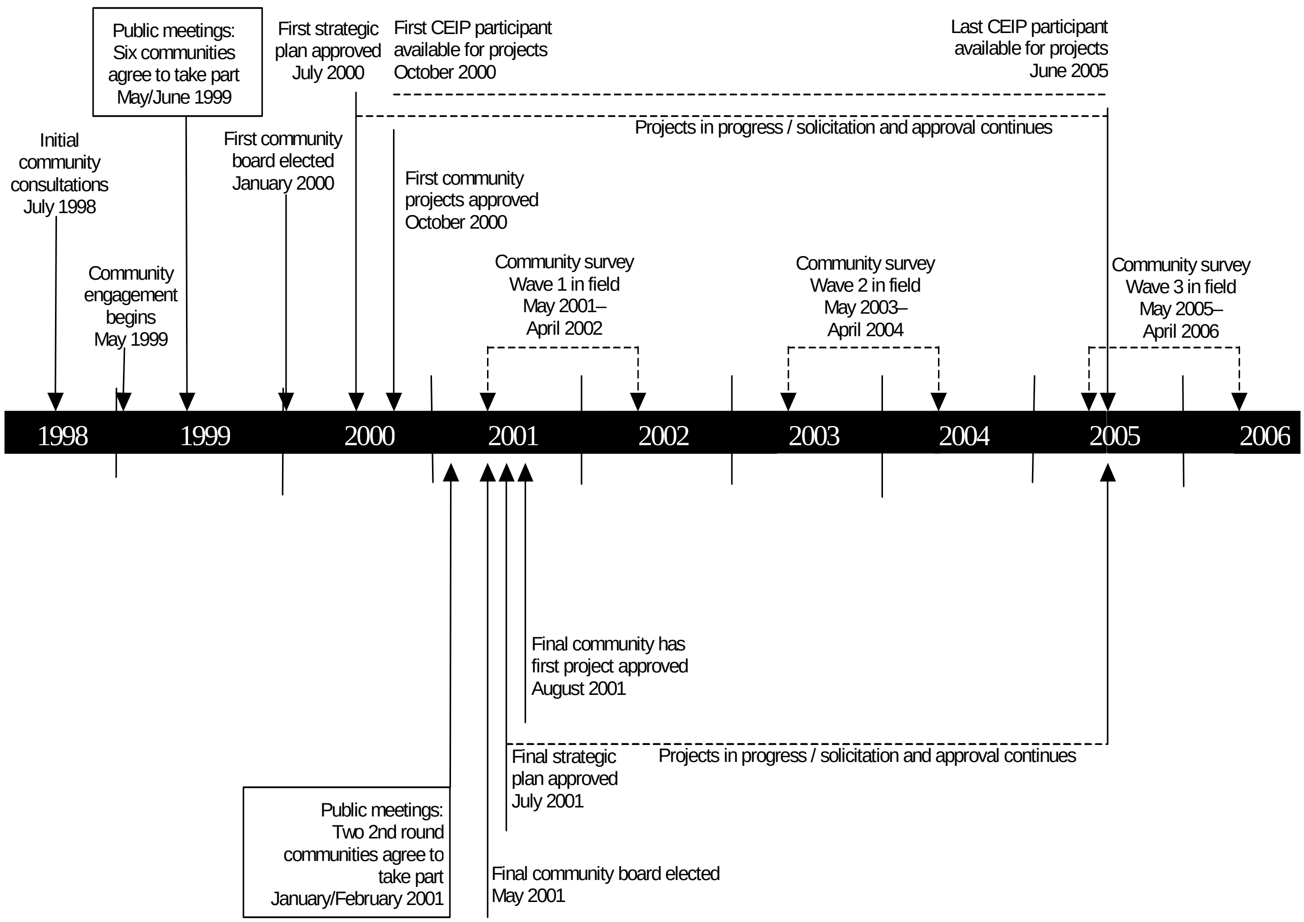


Figure A.1: CEIP Timeline

Table A.1: Summary Statistics

Variable	Mean	St. Dev.	Min.	Max.	Mean Comp. Com	Mean Prog. Com
Age (years)	48.3	(16.702)	14	94	48.25	48.36
Gender: male (dummy)	0.416	(0.493)	0	1	0.421	0.413
Married or living with a partner	0.570	(0.495)	0	1	0.587	0.557
Single, without children	0.318	(0.322)	0	1	0.321	0.316
Single with children	0.112	(0.315)	0	1	0.091	0.126
Couples with or without children	0.569	(0.495)	0	1	0.587	0.557
Household, including respondent	2.683	(1.356)	1	10	2.63	2.72
Number of children under 18 in HH	0.642	(0.964)	0	4	0.64	0.64
Number of adults in HH	1.978	(0.829)	1	7	1.93	2.01
High school	0.561	(0.496)	0	1	0.567	0.557
Bachelor’s degree	0.092	(0.288)	0	1	0.103	0.084
University (undergrad/grad) edu	0.03	(0.172)	0	1	0.041	0.024
Voc/apprenticeship diploma	0.227	(0.419)	0	1	0.22	0.23
Household income for year (’000 $)	39.034	(37.202)	1	997	42.7	36.57
Personal income for year (’000 $)	24.02	(29.571)	1	900	26.49	22.38
Household income below LICO	0.398	(0.489)	0	1	0.36	0.42
Gross hourly wage ($)	12.862	(7.205)	0.02	116.67	12.96	12.79
Union member on job	0.363	(0.481)	0	1	0.33	0.38
If not union member, are wages covered by union contract	0.115	(0.319)	0	1	0.09	0.14
Hours worked (per week)	40.35	(15.53)	0	164	42.3	38.85
Receiving pension from work	0.252	(0.434)	0	1	0.24	0.26
Receiving EI/SA	0.239	(0.427)	0	1	0.257	0.226

Prog. Com: Program communities; Comp. Com: Comparison communities.

Table A.2: Summary Statistics for Social Network Measures

Variable	Mean	St. Dev.	Min.	Max.	Mean Comp. Com	Mean Prog. Com
Family members respondent sees and talks to	9.654	(9.577)	0	97	9.783	9.568
Friends respondent sees and talks to	9.524	(12.551)	0	97	9.692	9.411
Network size (*family* + *friends*)	19.088	(17.856)	0	194	19.38	18.89
Total bonding contacts	24.174	(27.475)	0	288	25.32	23.42
Total bridging contacts	4.527	(7.459)	0	96	4.989	4.222
Total contacts that can help (*bond* + *bridge*)	28.61	(31.996)	0	384	30.23	27.69
Has linking contact (dummy)	0.304	(0.46)	0	1	0.32	0.29

Prog. Com: Program communities; Comp. Com: Comparison communities.

Table A.3: Summary Statistics for Potential Instruments

Variable	Mean	(Std. Dev.)	Mean Comp. Com	Mean Prog. Com
Project-related				
Heard about CEIP	0.227	(0.419)	0.15	0.27
Presently involved with CEIP	0.024	(0.152)	0.012	0.031
Presently paid by CEIP	0.006	(0.075)	0.003	0.007
Paid CEIP participant	0.003	(0.052)	0.001	0.004
Worker paid directly by proj sponsor	0.002	(0.041)	0.0003	0.003
Paid employee of CEIP consortium	0.001	(0.029)	0.0009	0.0008
Other paid involvement	0.002	(0.045)	0.0019	0.0022
Non-paid involvement with CEIP	0.010	(0.097)	0.0043	0.013
CEIP Control group member	0.003	(0.052)	0.0012	0.013
CEIP volunteer awaiting RA outcome	0.001	(0.031)	0.003	0.004
Resp lives in CEIP Program community	0.607	(0.488)	0	1
Non Project-related				
Number of years at current address	18.477	(17.06)	17.6	19.05
Number of years living in community	33.813	(20.584)	31.04	35.65
Number of years living in CB	42.562	(18.414)	41.21	43.46
Mother born in region?	0.754	(0.431)	0.73	0.77
Father born in region?	0.741	(0.438)	0.73	0.75
Respondent born in Cape Breton?	0.857	(0.35)	0.819	0.882

Prog. Com: Program communities; Comp. Com: Comparison communities.

Table A.4: Variable Definitions

Variable	Definition
Dependent Variable	
hrswork2	Hours worked in all jobs (per week)
Independent Variables	
ager	Age of respondent
agesq	Age-squared
male	Dummy, 1 if male
Education - highest level achieved	
hschool	Dummy, 1 if completed high school
college	Dummy, 1 if completed college
univ	Dummy, 1 if some university education
kid1	Dummy, 1 if no children under 18 in HH; **reference group**
kids2	Dummy, 1 if 1 child
kids3	Dummy, 1 if 2 children
kids4	Dummy, 1 if 3 children
kids5	Dummy, 1 if 4 or more children
Single, no children	Family status: single, with no children
Single with children	Family status: single with children
Couples with or without children	**reference group**
mar	Dummy, 1 if married or living with a partner
pension	Dummy, 1 if received pension from work or CPP/OAS/GIS in last year
welfare	Dummy, 1 if received EI/SA in last year
occ(1-10)	Dummy, 1 if occupation is one of 10 NOC groups
ind(1-10)	Dummy, 1 if industry of main/last job is one of 10 industries
town(1-13)	Dummy, 1 if living in one of 13 sample towns
union	Dummy, 1 if union member or wages covered by union contract
treat4	Paid CEIP participant (randomly assigned to Program group)
Social Network Measures	
link	Has linking contact? access to non-relative lawyer
bridge	Total bridging contacts
bond	Total bonding contacts
nwk	Network size (sum of family and friends)
tothelp	Total contacts that can help (sum of *bond* and *bridge*)
Instrumental Variables	
treat2	Presently involved with CEIP in any capacity
treat11	Respondent lives in CEIP program community
community2	Number of years living in community
community4	Respondent born in Cape Breton region
community5	Mother born in region

Table A.5: Test Statistics Definitions

Statistic	Definition
Model summary statistics	
R^2	R-squared[a]
F	F statistic
rmse	Root Mean Square Error
ll	Log likelihood
Tests of underidentification	
idstat	Kleibergen-Paap LM test statistic for underidentification
idp	p-value of underidentification LM statistic
Tests of weak identification	
widstat	Kleibergen-Paap F statistic for weak identification
cd	Shea's R-squared
cdf	Cragg-Donald F-statistic
Tests of weak-instrument robust inference	
arf	Anderson-Rubin F-test of significance of endogenous regressors
arfp	p-value of Anderson-Rubin F-test of endogenous regressors
archi2	Anderson-Rubin chi-sq test of significance of endogenous regressors
archi2p	p-value of Anderson-Rubin chi-sq test of endogenous regressors
sstat	Stock-Wright S statistic for significance of endogenous regressors
sstatp	p-value of Stock-Wright S statistic
Test of overidentification	
Hansen J statistic	Hansen J statistic of overidentification
jp	p-value of Hansen J statistic

[a]The R^2 has no statistical meaning in the context of GMM/IV estimation. The Explained Sum of Squares (ESS) under GMM is calculated using the actual values of regressors, not the instruments. The model's residuals are thus computed over a set of regressors different from those used to fit the model. When Residual Sum of Squares (RSS) exceeds Total Sum of Squares (TSS), the ESS and the R^2 are negative. A negative R^2 does not imply our parameter estimates are no good. Stata (2005) demonstrates results of simulations where the parameter estimates from IV estimation are quite good while the ESS is negative.

Table A.6: Effect of Linking Contacts (*link*) on Employment (*hours worked*)

VARIABLES	IV1	IV2	IV3
link	**17.30****	**17.93****	**18.87****
	(7.579)	**(7.758)**	**(8.249)**
ager	0.885***	0.884***	0.866***
	(0.213)	(0.216)	(0.226)
agesq	-0.0101***	-0.0101***	-0.00984***
	(0.00241)	(0.00244)	(0.00254)
male	4.161***	4.080***	4.024***
	(0.823)	(0.842)	(0.870)
hschool	-0.534	-0.504	-0.516
	(1.034)	(1.052)	(1.067)
college	-1.494	-1.488	-1.504
	(1.537)	(1.565)	(1.618)
univ	1.144	1.124	1.086
	(2.281)	(2.332)	(2.395)
kids2	-0.801	-0.811	-0.836
	(0.876)	(0.894)	(0.913)
kids3	-1.525*	-1.567*	-1.585*
	(0.905)	(0.925)	(0.948)
kids4	-3.028**	-3.082**	-2.993**
	(1.278)	(1.300)	(1.328)
Single, no children	-0.0887	-0.1000	-0.000117
	(0.811)	(0.827)	(0.841)
Single with children	-0.920	-0.884	-0.844
	(0.917)	(0.934)	(0.949)
treat11		-3.965	-3.357
		(4.479)	(4.490)
pension	-2.531*	-2.443*	-2.520*
	(1.296)	(1.327)	(1.346)
welfare	1.130*	1.202*	1.195*
	(0.649)	(0.666)	(0.679)
Constant	15.42***	19.15***	18.33***
	(3.757)	(5.702)	(5.733)

Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1
There are unreported additional coefficients.

Table A.7: Statistical Tests on the Effect of Linking Contacts (*link*) on Employment (*hours worked*)

VARIABLES	IV1	IV2	IV3
Observations	3775	3775	3791
R^2	0.0361	0.0182	-0.00908
F	17.20	16.24	15.89
rmse	15.27	15.42	15.63
ll	-15648	-15683	-15802
Tests of Identification			
idstat	15.30	15.07	13.50
idp	0.00915	0.00456	0.00117
widstat	3.122	3.847	6.778
cd	0.00459	0.00448	0.00401
cdf	3.420	4.172	7.508
arf	1.831	2.041	3.668
arfp	0.103	0.0859	0.0256
archi2	9.275	8.273	7.428
archi2p	0.0986	0.0821	0.0244
sstat	8.903	7.974	7.206
sstatp	0.113	0.0925	0.0272
Hansen J-statistic	2.437	1.591	0.830
jp	0.656	0.661	0.362

Table A.8: Effect of Bridging Contacts (*bridge*) on Employment (*hours worked*)

VARIABLES	IV1	IV2	IV3
bridge	**0.544***	**0.517**	**0.902****
	(0.319)	**(0.321)**	**(0.449)**
ager	1.213***	1.221***	1.179***
	(0.173)	(0.173)	(0.188)
agesq	-0.0136***	-0.0137***	-0.0132***
	(0.00216)	(0.00216)	(0.00233)
male	4.820***	4.835***	4.299***
	(0.737)	(0.735)	(0.878)
hschool	0.894	0.937	1.006
	(0.819)	(0.820)	(0.864)
college	0.950	1.013	1.093
	(1.054)	(1.057)	(1.109)
univ	3.797**	3.888**	3.734**
	(1.604)	(1.601)	(1.819)
kids2	-0.459	-0.442	-0.458
	(0.792)	(0.790)	(0.844)
kids3	-0.694	-0.713	-0.522
	(0.802)	(0.800)	(0.862)
kids4	-2.191*	-2.227*	-1.739
	(1.190)	(1.189)	(1.262)
Single, no children	0.0950	0.0939	0.316
	(0.782)	(0.780)	(0.838)
Single with children	-0.989	-0.998	-0.704
	(0.872)	(0.870)	(0.930)
treat11		-1.973	-0.726
		(3.107)	(3.187)
pension	-2.998**	-2.922**	-3.151**
	(1.255)	(1.256)	(1.379)
welfare	1.014	1.046*	1.122*
	(0.622)	(0.622)	(0.656)
Constant	12.05***	13.98***	11.58**
	(3.974)	(5.016)	(5.297)

Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

There are unreported additional coefficients.

Table A.9: Statistical Tests on the Effect of Bridging Contacts (*bridge*) on Employment (*hours worked*)

VARIABLES	IV1	IV2	IV3
Observations	3592	3592	3607
R^2	0.223	0.228	0.140
F	19.18	18.81	17.57
rmse	13.70	13.66	14.43
ll	-14499	-14488	-14745
Tests of Identification			
idstat	36.44	34.13	14.85
idp	7.76e-07	7.01e-07	0.000596
widstat	7.603	8.721	7.493
cd	0.0124	0.0121	0.00749
cdf	8.763	10.70	13.33
arf	1.810	2.053	3.673
arfp	0.107	0.0844	0.0255
archi2	9.174	8.324	7.443
archi2p	0.102	0.0804	0.0242
sstat	8.814	8.015	7.200
sstatp	0.117	0.0910	0.0273
Hansen J-statistic	5.175	4.812	0.816
jp	0.270	0.186	0.366

Table A.10: Effect of Bonding Contacts (*bond*) on Employment (*hours worked*)

VARIABLES	IV1	IV2	IV3
bond	**0.0793**	**0.0444**	**0.216**
	(0.127)	**(0.132)**	**(0.183)**
ager	1.143***	1.145***	1.171***
	(0.166)	(0.164)	(0.178)
agesq	-0.0125***	-0.0125***	-0.0129***
	(0.00210)	(0.00208)	(0.00228)
male	4.933***	5.071***	4.268***
	(0.846)	(0.853)	(1.048)
hschool	0.389	0.504	0.372
	(0.842)	(0.845)	(0.890)
college	0.191	0.378	0.0639
	(1.128)	(1.139)	(1.239)
univ	3.876**	4.168**	3.369*
	(1.672)	(1.686)	(1.864)
kids2	0.135	0.243	-0.117
	(0.831)	(0.833)	(0.916)
kids3	-0.495	-0.425	-0.703
	(0.850)	(0.848)	(0.929)
kids4	-1.535	-1.648	-1.048
	(1.223)	(1.228)	(1.321)
Single, no children	0.00442	0.0299	0.0232
	(0.784)	(0.784)	(0.822)
Single with children	-1.527*	-1.586*	-1.273
	(0.862)	(0.863)	(0.912)
treat11		-3.081	-1.908
		(3.078)	(3.377)
pension	-2.390*	-2.188*	-2.921*
	(1.305)	(1.301)	(1.596)
welfare	0.999	1.000	1.188*
	(0.632)	(0.631)	(0.680)
Constant	14.21***	18.22***	11.28
	(5.512)	(6.832)	(8.584)

Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

There are unreported additional coefficients.

Table A.11: Statistical Tests on the Effect of Bonding Contacts (*bond*) on Employment (*hours worked*)

VARIABLES	IV1	IV2	IV3
Observations	3539	3539	3550
R^2	0.234	0.243	0.129
F	18.62	18.31	16.63
rmse	13.40	13.32	14.29
ll	-14207	-14185	-14479
Tests of Identification			
idstat	23.67	21.64	12.55
idp	0.000251	0.000237	0.00188
widstat	4.858	5.457	6.277
cd	0.00570	0.00536	0.00295
cdf	3.975	4.678	5.158
arf	1.598	1.672	2.900
arfp	0.157	0.153	0.0552
archi2	8.100	6.784	5.878
archi2p	0.151	0.148	0.0529
sstat	7.817	6.580	5.715
sstatp	0.167	0.160	0.0574
j	7.272	6.403	3.425
jp	0.122	0.0936	0.0642

Table A.12: Effect of Network Size (*nwk*) on Employment (*hours worked*)

VARIABLES	IV1	IV2	IV3
nwk	**0.298****	**0.293***	**0.477****
	(0.152)	**(0.152)**	**(0.214)**
ager	1.365***	1.368***	1.442***
	(0.186)	(0.186)	(0.206)
agesq	-0.0150***	-0.0150***	-0.0157***
	(0.00226)	(0.00226)	(0.00249)
male	4.255***	4.264***	3.633***
	(0.813)	(0.811)	(0.993)
hschool	0.274	0.328	0.305
	(0.875)	(0.875)	(0.961)
college	0.276	0.337	0.303
	(1.119)	(1.119)	(1.238)
univ	3.606**	3.679**	3.500**
	(1.593)	(1.592)	(1.755)
kids2	0.134	0.115	0.391
	(0.807)	(0.805)	(0.901)
kids3	-0.750	-0.774	-0.623
	(0.797)	(0.796)	(0.879)
kids4	-1.816	-1.853	-1.294
	(1.208)	(1.206)	(1.334)
Single, no children	0.327	0.313	0.683
	(0.812)	(0.811)	(0.903)
Single with children	-1.149	-1.140	-0.972
	(0.876)	(0.875)	(0.942)
treat11		-2.776	-1.948
		(3.434)	(4.083)
	(1.241)	(1.242)	(1.396)
welfare	1.189*	1.218*	1.325*
	(0.639)	(0.639)	(0.696)
Constant	4.426	7.262	0.280
	(6.468)	(7.367)	(9.407)

Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

There are unreported additional coefficients.

Table A.13: Statistical Tests on the Effect of Network Size (*nwk*) on Employment (*hours worked*)

VARIABLES	IV1	IV2	IV3
Observations	3700	3700	3713
R^2	0.167	0.171	0.0225
F	17.41	17.09	14.48
rmse	14.16	14.14	15.34
ll	-15058	-15051	-15408
Tests of Identification			
idstat	21.98	21.73	11.81
idp	0.000527	0.000227	0.00273
widstat	4.559	5.633	6.089
cd	0.00920	0.00918	0.00566
cdf	6.717	8.382	10.38
arf	1.815	1.987	3.723
arfp	0.106	0.0937	0.0243
archi2	9.199	8.055	7.542
archi2p	0.101	0.0896	0.0230
sstat	8.837	7.777	7.318
sstatp	0.116	0.100	0.0258
j	3.987	3.355	0.00169
jp	0.408	0.340	0.967

Table A.14: Correlations between Social Network Measures

	nwk	numfam	numfri	tothelp	help1	help2	help3	help4
nwk	1.00							
numfam	0.92	1.00						
numfri	0.91	0.69	1.00					
tothelp	0.64	0.79	0.37	1.00				
help1	0.74	0.85	0.51	0.82	1.00			
help2	0.63	0.75	0.41	0.89	0.58	1.00		
help3	0.52	0.63	0.32	0.93	0.61	0.90	1.00	
help4	0.34	0.51	0.10	0.68	0.75	0.31	0.47	1.00

Table A.15: Effect of Linking Contacts on Hours Worked by Union Coverage

VARIABLES	IV1(U)	IV1(NU)	IV2(U)	IV2(NU)	IV3(U)	IV3(NU)
link	**-7.094**	**13.25****	**-9.853**	**14.40****	**-2.452**	**16.21****
	(11.40)	**(6.445)**	**(13.30)**	**(6.693)**	**(21.96)**	**(7.104)**
ager	0.506	1.062***	0.533	1.059***	0.459	1.058***
	(0.317)	(0.231)	(0.331)	(0.235)	(0.410)	(0.241)
agesq	-0.00612*	-0.0115***	-0.00643*	-0.0115***	-0.00556	-0.0115***
	(0.00369)	(0.00271)	(0.00385)	(0.00274)	(0.00470)	(0.00280)
male	5.852***	4.738***	6.048***	4.623***	5.558***	4.525***
	(1.361)	(0.880)	(1.464)	(0.913)	(1.970)	(0.936)
hschool	0.629	-0.0578	0.697	-0.0332	0.428	-0.149
	(1.271)	(1.225)	(1.301)	(1.257)	(1.457)	(1.269)
college	1.301	-0.530	1.681	-0.521	0.828	-0.659
	(2.247)	(1.655)	(2.463)	(1.699)	(3.495)	(1.724)
univ	7.251**	1.538	7.847**	1.348	6.367	1.136
	(3.066)	(2.783)	(3.433)	(2.891)	(4.923)	(2.928)
Single, no child	1.847*	-1.261	1.885*	-1.306	2.029**	-1.148
	(1.031)	(1.094)	(1.058)	(1.125)	(0.968)	(1.151)
Single w children	-1.380	-1.722	-1.537	-1.646	-1.171	-1.410
	(1.533)	(1.085)	(1.602)	(1.121)	(1.552)	(1.156)
pension	-0.119	-5.185***	-0.0814	-5.122***	0.000693	-4.969***
	(1.203)	(1.848)	(1.228)	(1.893)	(1.194)	(1.925)
welfare	-0.385	0.401	-0.546	0.486	0.0167	0.446
	(1.167)	(0.817)	(1.249)	(0.840)	(1.717)	(0.855)
Constant	29.86***	14.09***	32.12***	19.72***	29.48***	19.18***
	(6.694)	(4.636)	(8.469)	(7.076)	(8.293)	(7.112)
Observations	1515	2260	1515	2260	1524	2267
R^2	0.0677	0.204	-0.0166	0.185	0.164	0.149
F	6.153	14.29	5.746	13.27	6.998	12.89
rmse	12.22	15.35	12.76	15.53	11.63	15.86
ll	-5942	-9379	-6008	-9406	-5902	-9482

U - Union; NU - without Union Robust se's in parentheses. *** p<0.01, ** p<0.05, * p<0.1
There are unreported additional coefficients.

Table A.16: Effect of Bridging Contacts on Hours Worked by Union Coverage

VARIABLES	IV1(U)	IV1(NU)	IV2(U)	IV2(NU)	IV3(U)	IV3(NU)
bridge	**0.529**	**0.677***	**0.556**	**0.633**	**0.611**	**0.975***
	(0.546)	**(0.394)**	**(0.555)**	**(0.394)**	**(0.674)**	**(0.546)**
ager	0.499*	1.260***	0.506*	1.277***	0.521*	1.235***
	(0.271)	(0.220)	(0.272)	(0.219)	(0.282)	(0.239)
agesq	-0.00658**	-0.0132***	-0.00667**	-0.0135***	-0.00685**	-0.0129***
	(0.00314)	(0.00280)	(0.00316)	(0.00279)	(0.00325)	(0.00301)
male	4.835***	4.796***	4.850***	4.841***	4.885***	4.212***
	(1.063)	(0.992)	(1.066)	(0.987)	(1.110)	(1.217)
hschool	0.524	1.322	0.545	1.406	0.563	1.410
	(1.204)	(1.071)	(1.207)	(1.070)	(1.224)	(1.118)
college	0.557	1.170	0.551	1.295	0.758	1.249
	(1.563)	(1.393)	(1.565)	(1.395)	(1.584)	(1.453)
univ	4.747***	3.742	4.682**	3.793	4.631**	3.605
	(1.821)	(2.718)	(1.835)	(2.673)	(1.885)	(3.039)
Single, no child	1.719	-0.718	1.658	-0.764	1.782	-0.178
	(1.061)	(1.138)	(1.079)	(1.133)	(1.115)	(1.255)
Single w children	-0.229	-1.862*	-0.183	-1.860*	-0.107	-1.593
	(1.488)	(1.090)	(1.496)	(1.085)	(1.522)	(1.145)
pension	0.474	-6.422***	0.440	-6.308***	0.642	-6.940***
	(1.179)	(2.115)	(1.186)	(2.097)	(1.231)	(2.366)
welfare	0.569	0.592	0.552	0.653	0.649	0.566
	(0.889)	(0.851)	(0.891)	(0.849)	(0.914)	(0.886)
Constant	26.78***	9.505*	25.42***	12.80**	24.31***	10.50
	(6.701)	(4.947)	(7.831)	(6.269)	(7.968)	(6.528)
Observations	1444	2148	1444	2148	1452	2155
R^2	0.123	0.270	0.116	0.277	0.102	0.206
F	6.337	14.96	6.180	14.73	6.172	13.96
rmse	11.68	14.77	11.73	14.69	11.88	15.38
ll	-5598	-8831	-5604	-8820	-5654	-8947

U - Union; NU - without Union Robust se's in parentheses. *** p<0.01, ** p<0.05, * p<0.1
There are unreported additional coefficients.

Table A.17: Effect of Network Size on Hours Worked by Union Coverage

VARIABLES	IV1(U)	IV1(NU)	IV2(U)	IV2(NU)	IV3(U)	IV3(NU)
nwk	**0.180**	**0.367****	**0.184**	**0.368****	**0.241**	**0.606****
	(0.220)	**(0.182)**	**(0.232)**	**(0.182)**	**(0.256)**	**(0.308)**
ager	0.542*	1.455***	0.547*	1.464***	0.578*	1.537***
	(0.312)	(0.222)	(0.325)	(0.222)	(0.339)	(0.254)
agesq	-0.00641*	-0.0154***	-0.00646*	-0.0155***	-0.00676*	-0.0161***
	(0.00355)	(0.00278)	(0.00369)	(0.00279)	(0.00385)	(0.00316)
male	5.194***	3.941***	5.191***	3.935***	5.150***	2.845*
	(1.021)	(1.106)	(1.022)	(1.107)	(1.040)	(1.610)
hschool	-0.0312	0.708	-0.0317	0.800	-0.179	0.870
	(1.233)	(1.134)	(1.233)	(1.138)	(1.279)	(1.326)
college	-0.263	0.699	-0.273	0.780	-0.329	0.857
	(1.638)	(1.454)	(1.649)	(1.458)	(1.724)	(1.721)
univ	5.067**	3.005	5.046**	3.004	4.833**	2.987
	(2.096)	(2.295)	(2.133)	(2.294)	(2.239)	(2.554)
Single, no child	2.051**	-0.605	2.050*	-0.641	2.192**	0.0465
	(1.046)	(1.157)	(1.047)	(1.160)	(1.068)	(1.406)
Single w children	-0.379	-2.294**	-0.360	-2.239**	-0.348	-2.144*
	(1.568)	(1.098)	(1.605)	(1.102)	(1.645)	(1.246)
pension	-0.322	-5.936***	-0.328	-5.914***	-0.141	-6.615***
	(1.220)	(1.920)	(1.226)	(1.924)	(1.282)	(2.331)
welfare	0.786	0.549	0.796	0.608	0.945	0.420
	(1.123)	(0.846)	(1.139)	(0.849)	(1.181)	(0.961)
Constant	22.33**	0.125	21.95*	4.575	19.45	-4.134
	(10.85)	(7.580)	(12.87)	(8.787)	(13.75)	(12.51)
Observations	1486	2214	1486	2214	1493	2220
R^2	0.138	0.193	0.136	0.193	0.104	-0.0202
F	6.485	13.16	6.330	12.88	6.096	10.18
rmse	11.74	15.42	11.76	15.42	12.01	17.32
ll	-5769	-9199	-5770	-9199	-5829	-9481

U - Union; NU - without Union Robust se's in parentheses. *** p<0.01, ** p<0.05, * p<0.1
There are unreported additional coefficients.

Table A.18: Effect of Linking Contacts on Hours Worked by Union Coverage and Age

VARIABLES	Union below 30	Non-Union below 30	Union above 30	Non-Union above 30
link	**-80.97**	**16.02****	**-18.68**	**18.12***
	(72.68)	**(7.689)**	**(25.68)**	**(9.661)**
ager	34.37	7.385**	0.318	0.0970
	(36.17)	(3.692)	(0.607)	(0.346)
agesq	-0.691	-0.133*	-0.00380	-0.00143
	(0.751)	(0.0786)	(0.00682)	(0.00367)
male	17.93	1.500	6.400***	6.466***
	(14.45)	(1.480)	(2.106)	(1.179)
hschool	6.702	5.257**	1.107	-2.048
	(14.23)	(2.375)	(2.000)	(1.660)
college	27.30	1.886	3.672	-1.668
	(24.99)	(2.752)	(4.539)	(2.231)
univ	28.91	1.443	11.03	0.174
	(26.02)	(5.911)	(6.706)	(3.354)
Single, no child	-1.365	-1.222	1.924	-0.765
	(11.03)	(2.385)	(1.365)	(1.367)
Single w children	-11.89	-0.443	0.739	0.774
	(12.40)	(2.418)	(2.067)	(1.286)
pension	-5.132	9.660**	-0.146	-7.977***
	(16.74)	(4.696)	(1.827)	(2.042)
welfare	-9.767	-2.237	-1.721	0.629
	(11.43)	(1.608)	(1.981)	(1.056)
Constant	-365.5	-59.07	31.37**	36.55***
	(421.9)	(42.55)	(15.45)	(10.06)
Observations	190	584	1334	1683
R^2	-3.937	0.147	-0.486	0.086
F	0.521	5.802	3.142	8.194
rmse	37.87	14.44	14.45	16.26
ll	-960.1	-2388	-5456	-7082

Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

There are unreported additional coefficients.

Table A.19: Effect of Linking Contacts on Hours Worked by Age

VARIABLES	IV1(>30)	IV1(<30)	IV2(>30)	IV2(<30)	IV3(>30)	IV3(<30)
link	**12.54**	**29.43****	**12.52**	**33.19****	**14.00**	**33.76****
	(9.086)	**(12.27)**	**(9.090)**	**(14.33)**	**(10.27)**	**(14.91)**
ager	0.246	2.727	0.237	2.362	0.229	2.391
	(0.274)	(3.870)	(0.275)	(4.287)	(0.282)	(4.332)
agesq	-0.00298	-0.0376	-0.00289	-0.0309	-0.00277	-0.0319
	(0.00297)	(0.0807)	(0.00297)	(0.0892)	(0.00305)	(0.0901)
male	5.760***	1.818	5.780***	1.489	5.677***	1.464
	(0.907)	(1.662)	(0.909)	(1.882)	(0.979)	(1.900)
hschool	-1.400	6.341**	-1.424	7.061**	-1.509	7.108**
	(1.211)	(2.645)	(1.213)	(2.896)	(1.277)	(2.935)
college	-1.461	1.686	-1.476	2.267	-1.563	2.267
	(1.858)	(3.033)	(1.859)	(3.340)	(2.015)	(3.364)
univ	1.577	3.624	1.538	4.966	1.317	5.638
	(2.626)	(6.204)	(2.630)	(7.251)	(2.851)	(7.155)
kids2	-0.262	-4.144	-0.237	-4.444	-0.283	-4.395
	(0.872)	(2.892)	(0.874)	(3.334)	(0.898)	(3.328)
kids3	-1.282	-3.133	-1.292	-2.992	-1.305	-2.967
	(0.917)	(2.845)	(0.918)	(3.171)	(0.954)	(3.185)
kids4	-1.954	-10.08**	-1.970	-9.844*	-1.894	-9.752*
	(1.242)	(4.700)	(1.243)	(5.176)	(1.297)	(5.224)
Single, no child	0.164	-1.672	0.167	-1.442	0.305	-1.436
	(0.852)	(2.152)	(0.853)	(2.427)	(0.872)	(2.451)
Single w children	1.339	-0.136	1.328	0.256	1.313	0.176
	(0.978)	(2.558)	(0.979)	(2.931)	(0.996)	(2.928)
pension	-4.883***	11.94***	-4.888***	14.24***	-4.968***	14.31***
	(1.245)	(4.362)	(1.246)	(5.201)	(1.275)	(5.278)
Observations	3003	772	3003	772	3017	774
R^2	0.123	-0.200	0.123	-0.354	0.0905	-0.381
F	13.68	5.228	13.35	4.106	12.96	4.036
rmse	14.04	17.80	14.04	18.91	14.31	19.08
ll	-12195	-3318	-12194	-3365	-12309	-3380

>, <**30** Age greater/less than 30 years. Robust se's in parentheses. *** p<0.01, ** p<0.05, * p<0.1
There are unreported additional coefficients.

Table A.20: Effect of Linking Contacts on Hours Worked by Gender

VARIABLES	IV1(M)	IV1(F)	IV2(M)	IV2(F)	IV3(M)	IV3(F)
link	**5.969**	**17.74**	**6.648**	**18.04**	**6.876**	**24.33***
	(7.417)	**(11.98)**	**(7.454)**	**(12.02)**	**(8.806)**	**(14.75)**
ager	1.042***	0.798**	1.055***	0.791**	1.056***	0.662*
	(0.264)	(0.335)	(0.265)	(0.336)	(0.292)	(0.389)
agesq	-0.0110***	-0.00944**	-0.0111***	-0.00939**	-0.0111***	-0.00805*
	(0.00302)	(0.00384)	(0.00304)	(0.00384)	(0.00330)	(0.00437)
hschool	1.459	-2.235	1.498	-2.277	1.507	-2.938
	(1.114)	(1.952)	(1.119)	(1.957)	(1.122)	(2.276)
college	-0.416	-2.546	-0.440	-2.590	-0.202	-3.646
	(1.704)	(2.685)	(1.702)	(2.691)	(1.788)	(3.161)
univ	4.378*	-0.0297	4.242	-0.0638	4.292	-1.238
	(2.626)	(3.535)	(2.644)	(3.543)	(2.733)	(4.115)
kids2	1.396	-1.770*	1.343	-1.775*	1.356	-1.903
	(1.367)	(1.033)	(1.376)	(1.035)	(1.403)	(1.162)
kids3	0.654	-3.197***	0.545	-3.212***	0.506	-3.446**
	(1.233)	(1.188)	(1.246)	(1.191)	(1.256)	(1.369)
kids4	-1.915	-3.513**	-2.102	-3.527**	-1.793	-3.735**
	(1.956)	(1.622)	(1.968)	(1.624)	(1.998)	(1.823)
Single, no child	-0.291	-0.847	-0.321	-0.818	-0.199	-0.889
	(1.230)	(1.033)	(1.235)	(1.036)	(1.249)	(1.168)
Single w children	-6.391***	1.493	-6.187***	1.473	-6.134***	1.664
	(1.861)	(0.983)	(1.880)	(0.985)	(1.869)	(1.126)
pension	-5.229***	0.702	-5.232***	0.866	-5.338***	1.132
	(1.511)	(2.132)	(1.523)	(2.158)	(1.538)	(2.371)
welfare	2.162**	-0.155	2.230**	-0.102	2.236**	-0.0499
	(0.926)	(0.820)	(0.933)	(0.828)	(0.943)	(0.924)
Observations	1713	2062	1713	2062	1723	2068
R^2	0.225	-0.182	0.222	-0.195	0.221	-0.540
F	7.413	5.808	7.194	5.666	7.236	4.635
rmse	15.04	13.83	15.07	13.91	15.09	15.78
ll	-7074	-8343	-7078	-8355	-7121	-8639

M - Male; F - Female. Robust se's in parentheses. *** p<0.01, ** p<0.05, * p<0.1
There are unreported additional coefficients.

Table A.21: Effect of Bridging Contacts on Hours Worked by Gender

VARIABLES	IV1(M)	IV1(F)	IV2(M)	IV2(F)	IV3(M)	IV3(F)
bridge	**0.225**	**1.067***	**0.192**	**1.040***	**0.390**	**1.327****
	(0.365)	**(0.549)**	**(0.368)**	**(0.563)**	**(0.567)**	**(0.663)**
ager	1.161***	1.201***	1.186***	1.201***	1.153***	1.190***
	(0.241)	(0.258)	(0.243)	(0.258)	(0.265)	(0.264)
agesq	-0.0124***	-0.0138***	-0.0126***	-0.0139***	-0.0123***	-0.0137***
	(0.00286)	(0.00333)	(0.00287)	(0.00333)	(0.00306)	(0.00341)
hschool	1.993*	-0.684	2.048*	-0.674	2.103*	-0.614
	(1.074)	(1.165)	(1.075)	(1.166)	(1.088)	(1.195)
college	0.440	-0.101	0.491	-0.0825	0.769	-0.0574
	(1.524)	(1.403)	(1.524)	(1.405)	(1.550)	(1.435)
univ	5.048*	2.763	5.130**	2.812	4.793*	2.830
	(2.605)	(1.914)	(2.598)	(1.926)	(2.814)	(2.014)
kids2	1.287	-1.204	1.306	-1.198	1.187	-1.234
	(1.289)	(0.994)	(1.289)	(0.993)	(1.320)	(1.044)
kids3	1.007	-2.022*	0.949	-2.038*	0.959	-1.875
	(1.207)	(1.130)	(1.209)	(1.130)	(1.224)	(1.216)
kids4	-1.741	-2.150	-1.819	-2.185	-1.403	-2.063
	(1.825)	(1.638)	(1.828)	(1.645)	(1.874)	(1.703)
Single, no child	-0.269	0.697	-0.253	0.688	-0.289	0.880
	(1.334)	(1.053)	(1.337)	(1.053)	(1.363)	(1.099)
Single w children	-6.120***	1.849*	-6.010***	1.817*	-5.731***	1.939*
	(1.889)	(0.945)	(1.890)	(0.956)	(1.962)	(1.015)
pension	-5.095***	0.0961	-5.065***	0.164	-5.323***	0.270
	(1.559)	(1.990)	(1.556)	(2.011)	(1.646)	(2.052)
welfare	1.959**	0.271	1.995**	0.287	2.084**	0.213
	(0.940)	(0.807)	(0.940)	(0.809)	(0.961)	(0.848)
Observations	1637	1955	1637	1955	1646	1961
R^2	0.239	0.0110	0.240	0.0199	0.227	-0.0844
F	7.423	6.581	7.227	6.440	7.331	5.992
rmse	14.82	12.68	14.80	12.62	14.94	13.26
ll	-6736	-7739	-6734	-7731	-6786	-7851

M - Male; F - Female. Robust se's in parentheses. *** p<0.01, ** p<0.05, * p<0.1

There are unreported additional coefficients.

Table A.22: Effect of Network Size on Hours Worked by Gender

VARIABLES	IV1(M)	IV1(F)	IV2(M)	IV2(F)	IV3(M)	IV3(F)
nwk	**0.0775**	**0.717***	**0.0893**	**0.909***	**0.152**	**1.003***
	(0.141)	**(0.380)**	**(0.142)**	**(0.488)**	**(0.193)**	**(0.577)**
ager	1.173***	1.728***	1.203***	1.871***	1.207***	1.921***
	(0.237)	(0.404)	(0.239)	(0.474)	(0.243)	(0.537)
agesq	-0.0121***	-0.0200***	-0.0125***	-0.0216***	-0.0124***	-0.0222***
	(0.00281)	(0.00498)	(0.00283)	(0.00574)	(0.00287)	(0.00647)
hschool	1.473	-1.289	1.516	-1.695	1.493	-1.630
	(1.106)	(1.557)	(1.110)	(1.764)	(1.129)	(1.925)
college	-0.124	-0.632	-0.138	-1.080	0.00145	-1.107
	(1.571)	(1.832)	(1.572)	(2.050)	(1.605)	(2.227)
univ	4.897**	1.891	4.812**	1.154	4.845**	0.965
	(2.332)	(2.706)	(2.333)	(3.106)	(2.350)	(3.429)
kids2	2.256*	-1.679	2.259*	-1.663	2.468*	-1.658
	(1.268)	(1.174)	(1.272)	(1.259)	(1.354)	(1.362)
kids3	1.184	-2.717**	1.084	-2.767**	1.105	-2.638*
	(1.169)	(1.262)	(1.179)	(1.357)	(1.200)	(1.474)
kids4	-1.287	-1.715	-1.435	-1.343	-0.958	-1.117
	(1.833)	(1.911)	(1.839)	(2.089)	(1.898)	(2.306)
Single, no child	8.25e-05	0.312	-0.00542	0.513	0.175	0.667
	(1.272)	(1.232)	(1.273)	(1.352)	(1.341)	(1.498)
Single w children	-6.528***	1.773	-6.308***	1.932	-6.198***	1.926
	(1.905)	(1.110)	(1.919)	(1.208)	(1.950)	(1.321)
pension	-5.070***	0.667	-5.110***	0.578	-5.417***	0.652
	(1.545)	(2.174)	(1.546)	(2.287)	(1.654)	(2.460)
welfare	2.171**	0.00899	2.237**	0.0392	2.298**	-0.0248
	(0.954)	(0.922)	(0.958)	(0.988)	(0.986)	(1.065)
Observations	1667	2033	1667	2033	1674	2039
R^2	0.227	-0.349	0.227	-0.678	0.211	-0.869
F	7.035	4.596	6.871	4.019	6.750	3.521
rmse	15.02	14.79	15.02	16.50	15.17	17.39
ll	-6881	-8362	-6882	-8584	-6927	-8717

M - Male; F - Female. Robust se's in parentheses. *** p<0.01, ** p<0.05, * p<0.1

There are unreported additional coefficients.

Table A.23: Effect of Linking Contacts on Hours Worked by High School Completion

VARIABLES	IV1(HS)	IV1(NHS)	IV2(HS)	IV2(NHS)	IV3(HS)	IV3(NHS)
link	**12.26****	**-20.60**	**12.98****	**-7.061**	**19.94****	**4.896**
	(5.753)	**(13.90)**	**(5.900)**	**(16.37)**	**(7.810)**	**(19.70)**
ager	0.941***	1.962***	0.928***	1.748***	0.776***	1.635***
	(0.203)	(0.454)	(0.205)	(0.431)	(0.243)	(0.396)
agesq	-0.0111***	-0.0197***	-0.0110***	-0.0174***	-0.00944***	-0.0162***
	(0.00242)	(0.00541)	(0.00243)	(0.00507)	(0.00282)	(0.00459)
male	3.491***	8.369***	3.465***	6.923***	3.030***	5.703**
	(0.713)	(2.569)	(0.716)	(2.628)	(0.832)	(2.762)
kids2	-0.851	6.143**	-0.847	5.193*	-0.783	3.863
	(0.756)	(2.718)	(0.758)	(2.656)	(0.850)	(2.999)
kids3	-1.276*	4.731*	-1.303*	4.041*	-1.407	2.809
	(0.775)	(2.529)	(0.779)	(2.386)	(0.878)	(2.780)
kids4	-2.317*	2.927	-2.355**	0.499	-2.202*	-2.179
	(1.186)	(3.971)	(1.190)	(4.285)	(1.295)	(5.451)
Single, no child	0.433	2.211	0.458	2.239	0.626	1.934
	(0.771)	(2.101)	(0.774)	(1.931)	(0.858)	(1.960)
Single w children	-0.860	-4.746	-0.853	-3.998	-0.742	-2.833
	(0.832)	(3.369)	(0.834)	(3.152)	(0.920)	(3.381)
treat11			-1.114	-8.308	-1.532	-9.882
			(1.947)	(7.033)	(2.634)	(9.495)
pension	-2.035	-3.190	-1.925	-4.195	-1.784	-5.108*
	(1.310)	(2.702)	(1.328)	(2.577)	(1.474)	(2.915)
welfare	0.563	0.745	0.604	0.108	0.857	0.0294
	(0.616)	(1.691)	(0.621)	(1.603)	(0.700)	(1.529)
Observations	3619	627	3619	627	3631	632
R^2	0.116	0.082	0.100	0.302	-0.112	0.336
F	16.17	4.510	15.74	5.535	13.11	6.422
rmse	13.68	18.68	13.80	16.28	15.36	15.85
ll	-14602	-2725	-14635	-2639	-15070	-2643

HS/NHS: High School completed/drop out. Robust se's in parentheses. *** p<0.01, ** p<0.05, * p<0.1
There are unreported additional coefficients.

Table A.24: Effect of Bridging Contacts on Hours Worked by High School Completion

VARIABLES	IV1(HS)	IV1(NHS)	IV2(HS)	IV2(NHS)	IV3(HS)	IV3(NHS)
bridge	**0.560**	**0.364**	**0.604**	**0.339**	**0.929***	**-0.0380**
	(0.375)	**(0.365)**	**(0.404)**	**(0.359)**	**(0.517)**	**(0.453)**
ager	1.166***	1.705***	1.163***	1.735***	1.117***	1.841***
	(0.178)	(0.348)	(0.178)	(0.349)	(0.195)	(0.357)
agesq	-0.0136***	-0.0172***	-0.0135***	-0.0175***	-0.0130***	-0.0187***
	(0.00225)	(0.00418)	(0.00225)	(0.00419)	(0.00248)	(0.00428)
male	3.981***	4.859***	3.925***	4.822***	3.561***	5.644***
	(0.721)	(1.800)	(0.746)	(1.794)	(0.864)	(1.896)
kids2	-0.782	3.380	-0.791	3.283	-0.788	3.221
	(0.752)	(2.262)	(0.754)	(2.270)	(0.829)	(2.239)
kids3	-0.776	3.719	-0.750	3.539	-0.589	2.786
	(0.764)	(2.267)	(0.769)	(2.277)	(0.838)	(2.349)
kids4	-1.578	-0.677	-1.530	-0.833	-0.923	-1.806
	(1.177)	(3.609)	(1.188)	(3.588)	(1.266)	(3.581)
Single, no child	0.745	3.002	0.755	2.802	1.007	2.887
	(0.782)	(2.033)	(0.783)	(2.035)	(0.844)	(2.074)
Single w children	-0.727	-2.256	-0.681	-2.040	-0.386	-1.879
	(0.888)	(2.856)	(0.903)	(2.854)	(0.987)	(2.834)
treat11			0.452	-7.902	0.976	-7.700
			(1.499)	(7.776)	(1.620)	(7.012)
pension	-2.670**	-4.660**	-2.731**	-4.626**	-2.926**	-4.275*
	(1.283)	(2.307)	(1.300)	(2.293)	(1.417)	(2.293)
pension	-2.670**	-4.660**	-2.731**	-4.626**	-2.926**	-4.275*
	(1.283)	(2.307)	(1.300)	(2.293)	(1.417)	(2.293)
welfare	0.538	0.117	0.544	0.0863	0.738	0.295
	(0.618)	(1.519)	(0.618)	(1.516)	(0.663)	(1.512)
Observations	3469	585	3469	585	3480	590
R^2	0.188	0.347	0.181	0.350	0.099	0.344
F	17.04	6.142	16.65	6.028	15.28	5.907
rmse	13.10	15.74	13.17	15.70	13.82	15.73
ll	-13848	-2442	-13864	-2441	-14077	-2463

HS/NHS: High School completed/drop out. Robust se's in parentheses. *** p<0.01, ** p<0.05, * p<0.1
There are unreported additional coefficients.

Table A.25: Effect of Network Size on Hours Worked by High School Completion

VARIABLES	IV1(HS)	IV1(NHS)	IV2(HS)	IV2(NHS)	IV3(HS)	IV3(NHS)
nwk	**0.224***	**0.0701**	**0.348****	**0.223**	**0.542****	**0.0705**
	(0.122)	**(0.221)**	**(0.163)**	**(0.255)**	**(0.225)**	**(0.439)**
ager	1.282***	1.590***	1.347***	1.701***	1.423***	1.687***
	(0.179)	(0.346)	(0.190)	(0.363)	(0.216)	(0.367)
agesq	-0.0147***	-0.0154***	-0.0153***	-0.0165***	-0.0160***	-0.0165***
	(0.00221)	(0.00407)	(0.00230)	(0.00428)	(0.00260)	(0.00426)
male	3.769***	5.562**	3.443***	4.504*	2.968***	5.587
	(0.664)	(2.255)	(0.726)	(2.476)	(0.860)	(3.637)
kids2	-0.970	5.993**	-0.915	6.399**	-0.842	5.510*
	(0.748)	(2.373)	(0.762)	(2.570)	(0.861)	(2.840)
kids3	-1.358*	4.632**	-1.424*	4.936**	-1.548*	3.909
	(0.775)	(2.318)	(0.792)	(2.419)	(0.901)	(3.220)
kids4	-1.881	-0.212	-1.652	-1.374	-1.048	-1.325
	(1.209)	(3.319)	(1.240)	(3.351)	(1.383)	(3.530)
Single, no child	0.355	2.780	0.403	2.998	0.515	2.061
	(0.770)	(2.171)	(0.782)	(2.221)	(0.870)	(2.822)
Single w children	-0.904	-3.924	-0.712	-3.162	-0.477	-3.366
	(0.844)	(2.938)	(0.871)	(3.050)	(0.969)	(3.024)
pension	-2.958**	-4.524*	-3.349**	-4.696*	-4.034***	-4.389*
	(1.253)	(2.316)	(1.311)	(2.469)	(1.533)	(2.419)
welfare	0.695	-0.351	0.840	-0.382	1.129	-0.254
	(0.625)	(1.523)	(0.644)	(1.619)	(0.730)	(1.556)
Observations	3567	602	3567	602	3577	606
R^2	0.188	0.337	0.121	0.297	-0.058	0.340
F	15.58	6.066	14.83	5.054	12.22	5.560
rmse	13.07	15.93	13.60	16.40	14.93	15.84
ll	-14230	-2521	-14371	-2538	-14745	-2534

HS/NHS: High School completed/drop out. Robust se's in parentheses. *** p<0.01, ** p<0.05, * p<0.1
There are unreported additional coefficients.

Table A.26: Effect of Linking Contacts on Hours Worked by Skill Level

VARIABLES	IV1(HS)	IV1(LS)	IV2(HS)	IV2(LS)	IV3(HS)	IV3(LS)
link	**2.754**	**4.534**	**2.163**	**9.938**	**12.08***	**17.85**
	(5.378)	**(13.05)**	**(5.462)**	**(14.21)**	**(7.321)**	**(16.54)**
ager	0.998***	1.118***	1.011***	0.989**	0.858***	0.816*
	(0.186)	(0.389)	(0.187)	(0.407)	(0.209)	(0.459)
agesq	-0.0117***	-0.0124***	-0.0119***	-0.0110**	-0.0104***	-0.00907*
	(0.00224)	(0.00428)	(0.00225)	(0.00443)	(0.00246)	(0.00496)
male	4.065***	5.853***	4.050***	5.276***	3.854***	4.602**
	(0.711)	(1.528)	(0.711)	(1.652)	(0.759)	(1.865)
hschool	-2.033	1.944*	-1.953	1.739	-2.836*	1.497
	(1.391)	(1.101)	(1.398)	(1.173)	(1.463)	(1.294)
college	-2.849*		-2.732*		-4.024**	
	(1.577)		(1.588)		(1.735)	
univ	-0.853		-0.654		-2.569	
	(1.964)		(1.991)		(2.267)	
Single, no child	0.474	-0.624	0.460	-0.926	0.881	-0.991
	(0.895)	(1.180)	(0.896)	(1.257)	(0.954)	(1.363)
Single w children	-0.584	-1.680	-0.598	-1.483	-0.315	-1.277
	(0.980)	(1.118)	(0.981)	(1.145)	(1.046)	(1.231)
treat11			-1.444	-5.866	-0.144	-7.810
			(2.238)	(5.941)	(2.690)	(6.960)
pension	-3.435***	-1.299	-3.357***	-1.364	-3.191**	-1.805
	(1.223)	(1.567)	(1.230)	(1.655)	(1.281)	(1.814)
welfare	-0.308	1.506*	-0.333	1.622**	0.134	1.690*
	(0.650)	(0.787)	(0.652)	(0.821)	(0.715)	(0.888)
Constant	15.69**	13.51**	16.92**	20.60**	18.92**	23.39**
	(7.373)	(5.829)	(7.617)	(9.025)	(7.985)	(10.16)
Observations	2558	2162	2558	2162	2566	2171
R^2	0.211	0.282	0.209	0.231	0.131	0.0710
F	11.45	14.53	11.18	13.27	10.45	11.71
rmse	12.81	13.86	12.82	14.34	13.45	15.76
ll	-10154	-8752	-10156	-8826	-10311	-9068

HS - High Skilled; LS - Low Skilled. Robust se's in parentheses. *** p<0.01, ** p<0.05, * p<0.1
There are unreported additional coefficients.

Table A.27: Effect of Link, Bridge, Bond, Tothelp, Nwk on Hours Worked under OLS

VARIABLES	link	bridge	bond	tothelp	nwk
link	**2.095***** **(0.496)**				
bridge		**0.132***** **(0.0441)**			
bond			**0.0156** **(0.0111)**		
tothelp/nwk				**0.0213**** **(0.0102)**	**0.0369**** **(0.0159)**
ager	1.180*** (0.157)	1.247*** (0.168)	1.124*** (0.163)	1.160*** (0.170)	1.218*** (0.161)
agesq	-0.0130*** (0.00196)	-0.0137*** (0.00211)	-0.0121*** (0.00204)	-0.0126*** (0.00214)	-0.0133*** (0.00201)
male	5.039*** (0.620)	5.233*** (0.625)	5.113*** (0.637)	5.087*** (0.638)	5.009*** (0.628)
hschool	0.668 (0.794)	1.032 (0.804)	0.602 (0.820)	0.955 (0.826)	0.678 (0.810)
college	0.794 (1.016)	1.354 (1.033)	0.620 (1.045)	0.925 (1.051)	0.937 (1.033)
univ	4.648*** (1.461)	5.014*** (1.475)	4.835*** (1.507)	5.010*** (1.497)	4.919*** (1.482)
Single, no child	0.00903 (0.754)	0.247 (0.778)	0.186 (0.780)	0.324 (0.797)	-0.00183 (0.771)
Single w children	-1.397* (0.818)	-1.304 (0.834)	-1.711** (0.836)	-1.554* (0.842)	-1.519* (0.830)
pension	-2.747** (1.166)	-2.880** (1.213)	-2.180* (1.206)	-2.122* (1.252)	-2.668** (1.174)
Constant	16.76*** (4.669)	15.01*** (4.787)	24.97*** (5.664)	22.79*** (5.785)	23.97*** (5.542)

There are unreported additional coefficients.

Regression statistics are presented in the next table.

Table A.28: Regression Statistics on the Effect of Link, Bridge, Bond, Tothelp, Nwk on Hours Worked under OLS

VARIABLES	link	bridge	bond	tothelp	nwk
Observations	3818	3633	3575	3448	3737
R^2	0.254	0.258	0.247	0.253	0.246
F	19.97	19.28	18.40	17.97	18.82
rmse	13.52	13.48	13.38	13.33	13.54
ll	-15338	-14583	-14323	-13799	-15017

Robust se's in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table A.29: Effect of link, lnbridge, lnbond, lntothelp, lnnwk on lnhrswork

VARIABLES	link	lnbridge	lnbond	lntothelp	lnnwk
link	**0.0591*** **				
	(0.0161)				
lnbridge		**0.0327*** **			
		(0.0111)			
lnbond			**0.0311** **		
			(0.0125)		
lnhelp/lnnwk				**0.0349*** **	**0.029*** **
				(0.0131)	**(0.0111)**
ager	0.0514***	0.0571***	0.0499***	0.0516***	0.0527***
	(0.00644)	(0.00729)	(0.00659)	(0.00687)	(0.00664)
agesq	-0.000574***	-0.000639***	-0.000549***	-0.000572***	-0.000586***
	(7.87e-05)	(9.01e-05)	(8.09e-05)	(8.50e-05)	(8.13e-05)
male	0.117***	0.136***	0.120***	0.120***	0.116***
	(0.0228)	(0.0241)	(0.0234)	(0.0238)	(0.0230)
hschool	0.0585**	0.0673**	0.0513*	0.0594**	0.0625**
	(0.0267)	(0.0283)	(0.0272)	(0.0278)	(0.0274)
college	0.0612*	0.0747**	0.0498	0.0560	0.0681*
	(0.0365)	(0.0378)	(0.0373)	(0.0379)	(0.0371)
univ	0.162***	0.166***	0.154***	0.161***	0.173***
	(0.0476)	(0.0509)	(0.0487)	(0.0499)	(0.0483)
Single, no child	0.00612	0.0267	0.0142	0.0206	0.00436
	(0.0253)	(0.0267)	(0.0260)	(0.0267)	(0.0258)
Single w children	-0.0453	-0.0194	-0.0487	-0.0451	-0.0489
	(0.0314)	(0.0334)	(0.0318)	(0.0321)	(0.0318)
pension	-0.0923**	-0.0745	-0.0679	-0.0700	-0.0927**
	(0.0431)	(0.0464)	(0.0437)	(0.0458)	(0.0441)
Constant	2.553***	2.377***	2.514***	2.487***	2.459***
	(0.209)	(0.223)	(0.242)	(0.225)	(0.215)

There are unreported additional coefficients.

Regression statistics are presented in the next table.

Table A.30: Regression Statistics on the Effect of link, lnbridge, lnbond, lntothelp, lnnwk on lnhrswork

VARIABLES	link	lnbridge	lnbond	lntothelp	lnnwk
Observations	3818	3362	3563	3439	3731
R^2	0.209	0.218	0.211	0.213	0.203
F	15.95	15.08	14.86	14.52	15.02
rmse	0.442	0.438	0.435	0.436	0.444
ll	-2280	-1972	-2070	-2005	-2243

Robust se's in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table A.31: Heterogeneity of Impact of Social Networks (OLS)

VARIABLES	link	bridge	bond	nwk	tothelp
By Gender					
Male	**2.659*** **	**0.162*** **	**0.0162**	**0.0337**	**0.0272** **
	(0.808)	(0.0534)	(0.0148)	(0.0213)	(0.0134)
Female	**1.534*** **	**0.0271**	**0.0009**	**0.0246**	**-0.00208**
	(0.588)	(0.0633)	(0.0135)	(0.0224)	(0.0119)
By Union coverage					
Union	**2.203*** **	**0.0132**	**0.0121**	**0.0107**	
	(0.645)	(0.0507)	(0.0132)	(0.0184)	
No Union	**1.909*** **	**0.200*** **	**0.0148**	**0.0467** **	
	(0.694)	(0.0576)	(0.0150)	(0.0220)	
By Age					
Old	**1.965*** **	**0.121*** **		**0.0282***	
	(0.526)	(0.0431)		(0.0165)	
Young	**2.732** **	**0.217**		**0.0702**	
	(1.162)	(0.155)		(0.0437)	
By High School					
High School	**2.304*** **	**0.132*** **		**0.0657*** **	
	(0.578)	(0.0575)		(0.0191)	
No High School	**1.264**	**0.203** **		**-0.0268**	
	(0.940)	(0.0675)		(0.0267)	
By Skill Level					
High Skilled	**3.481*** **	**0.104** **		**0.0230**	
	(0.559)	(0.0449)		(0.0164)	
Low Skilled	**1.348** **	**0.160** **		**0.0446** **	
	(0.684)	(0.0638)		(0.0223)	

Standard errors, in parentheses. *** $p<0.01$, ** $p<0.05$, * $p<0.1$

Each of the result displayed above comes from a separate regression.

Table A.32: Heterogeneity of Impact of Social Networks (Log OLS)

VARIABLES	link	lnbridge	lnbond	lnnwk	lntothelp
By Gender					
Male	**0.0792*** **	**0.0471*** **	**0.0404** **	**0.0299***	**0.0484*** **
	(0.0237)	(0.0151)	(0.0177)	(0.0161)	(0.0184)
Female	**0.0402***	**0.00827**	**0.00481**	**0.0153**	**0.00418**
	(0.0218)	(0.0151)	(0.0165)	(0.0149)	(0.0171)
By Union coverage					
Union	**0.0710*** **	**0.00991**	**0.0220***	**0.00928**	
	(0.0182)	(0.0124)	(0.0128)	(0.0120)	
No Union	**0.0477** **	**0.0453*** **	**0.0336***	**0.0393** **	
	(0.0238)	(0.0161)	(0.0183)	(0.0161)	
By Age					
Old	**0.0531*** **	**0.0254*** **		**0.0282*** **	
	(0.0158)	(0.00955)		(0.0104)	
Young	**0.0867** **	**0.0708***		**0.0481**	
	(0.0439)	(0.0368)		(0.0346)	
By High School					
High School	**0.0648*** **	**0.0327*** **		**0.0419*** **	
	(0.0193)	(0.0137)		(0.0123)	
No High School	**0.0492**	**0.0404**		**0.0265**	
	(0.0296)	(0.0184)		(0.0207)	
By Skill Level					
High Skilled	**0.0941*** **	**0.0199***		**0.0171**	
	(0.0173)	(0.0118)		(0.0116)	
Low Skilled	**0.0349**	**0.0462*** **		**0.0419*** **	
	(0.0228)	(0.0158)		(0.0152)	

Standard errors, in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Each of the result displayed above comes from a separate regression.

A.0.1 STATA Output

We include here complete output from STATA for one model - the effect of linking contacts (*link*) on employment (*hours worked*). A summary of these results have been presented in column (3) of Table A.6.

The output displays the complete estimation results from first-stage and second-stage regressions. It also includes all the various tests of identification - the overidentification test, the weak and underidentification tests, the test of exogeneity, and tests of redundancy of instruments.

```
ivreg2 hrswork2 `control1' (link =`inst3') if survey==1&treat4==0
[pweight=hwtall], gmm2s robust first savefirst endog(link) orthog(treat2)

(7610 missing values generated)
(sum of wgt is     4.0643e+03)
Warning - collinearities detected
Vars dropped:       ind10

Stored estimation results
-------------------------
----------------------------------------------------------------------------
        name | command      depvar       npar  title
-------------+--------------------------------------------------------------
_ivreg2_link | ivreg2       link           47  First-stage regression: link
----------------------------------------------------------------------------

First-stage regressions
-----------------------

First-stage regression of link:

OLS estimation
--------------

Estimates efficient for homoskedasticity only
Statistics robust to heteroskedasticity

                                                      Number of obs =     3791
                                                      F( 46,  3744) =     5.52
                                                      Prob > F      =   0.0000
Total (centered) SS     =  914.2507489                Centered R2   =   0.0622
Total (uncentered) SS   =  1539.135494                Uncentered R2 =   0.4429
Residual SS             =  857.4173339                Root MSE      =    .4786

------------------------------------------------------------------------------
             |               Robust
        link |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
        ager |   .0186659   .0047256     3.95   0.000     .0094008     .027931
       agesq |  -.0001829   .0000589    -3.11   0.002    -.0002983   -.0000675
        male |   .0584944   .0212009     2.76   0.006     .0169279     .100061
     hschool |   .0718469   .0236995     3.03   0.002     .0253817    .1183121
     college |   .1373056   .0337591     4.07   0.000     .0711175    .2034937
        univ |   .2237973   .0502741     4.45   0.000     .1252301    .3223645
       kids2 |   .0329441   .0266274     1.24   0.216    -.0192615    .0851496
       kids3 |    .049361   .0270464     1.83   0.068    -.0036662    .1023882
       kids4 |   .0491545   .0385354     1.28   0.202     -.026398     .124707
       kids5 |   .0158985   .0723963     0.22   0.826    -.1260414    .1578385
  marstatus1 |  -.0007362   .0254482    -0.03   0.977    -.0506298    .0491574
  marstatus2 |  -.0329697   .0290657    -1.13   0.257    -.0899559    .0240165
     treat11 |   .0415755   .1310724     0.32   0.751    -.2154049    .2985558
        occ2 |  -.0578048   .0434859    -1.33   0.184    -.1430632    .0274536
        occ3 |  -.1123868   .0544382    -2.06   0.039    -.2191181   -.0056554
        occ4 |  -.0696816   .0507665    -1.37   0.170    -.1692143    .0298511
        occ5 |    .018254   .0489405     0.37   0.709    -.0776987    .1142066
        occ6 |  -.0934677   .0627973    -1.49   0.137    -.2165879    .0296525
        occ7 |  -.0691148   .0393613    -1.76   0.079    -.1462865    .0080568
        occ8 |  -.0799989   .0447272    -1.79   0.074    -.1676909    .0076932
        occ9 |  -.0858267   .0740249    -1.16   0.246    -.2309598    .0593064
       occ10 |  -.1610309   .0661428    -2.43   0.015    -.2907103   -.0313514
```

Figure A.2: Effect of Linking Contacts on Employment - Stata output page 1

```
        ind2 |  -.0639908    .082211    -0.78   0.436    -.2251734    .0971918
        ind3 |  -.0448847   .0514906    -0.87   0.383     -.145837    .0560677
        ind4 |    -.11745   .0581787    -2.02   0.044     -.231515   -.003385
        ind5 |  -.1128995   .0419976    -2.69   0.007      -.19524   -.030559
        ind6 |  -.0199237   .0428392    -0.47   0.642    -.1039143    .0640668
        ind7 |  -.1045695   .0446511    -2.34   0.019    -.1921124   -.0170265
        ind8 |  -.0669216   .0472797    -1.42   0.157    -.1596181    .0257748
        ind9 |   -.091025   .0522974    -1.74   0.082    -.1935591    .0115091
     pension |  -.0124319   .0382265    -0.33   0.745    -.0873788    .0625149
     welfare |  -.0253304   .0193163    -1.31   0.190    -.0632019    .0125411
       town2 |  -.0087951   .0411245    -0.21   0.831    -.0894238    .0718335
       town3 |   .1669177   .0374587     4.46   0.000     .0934762    .2403593
       town4 |    .082008   .1378495     0.59   0.552    -.1882594    .3522754
       town5 |   .1990102    .138103     1.44   0.150    -.0717542    .4697746
       town6 |   .1380109   .1379112     1.00   0.317    -.1323775    .4083993
       town7 |   .1034298   .0369064     2.80   0.005     .0310712    .1757884
       town8 |   .1037035   .1326813     0.78   0.434    -.1564311    .3638381
       town9 |   .0250891   .1416796     0.18   0.859    -.2526876    .3028658
      town10 |   .0765261   .0362064     2.11   0.035       .00554    .1475123
      town11 |   .1325377   .0377201     3.51   0.000     .0585838    .2064916
      town12 |   .1916055   .1352286     1.42   0.157    -.0735233    .4567343
      town13 |   .1101533   .1371661     0.80   0.422    -.1587741    .3790808
      treat2 |   .1234023   .0501942     2.46   0.014     .0249917     .221813
  community5 |   .0593432   .0209997     2.83   0.005     .0181713    .1005151
       _cons |  -.1700813   .1704402    -1.00   0.318     -.504246    .1640834
------------------------------------------------------------------------------
Included instruments: ager agesq male hschool college univ kids2 kids3 kids4
                      kids5 marstatus1 marstatus2 treat11 occ2 occ3 occ4 occ5
                      occ6 occ7 occ8 occ9 occ10 ind2 ind3 ind4 ind5 ind6 ind7
                      ind8 ind9 pension welfare town2 town3 town4 town5 town6
                      town7 town8 town9 town10 town11 town12 town13 treat2
                      community5
------------------------------------------------------------------------------
Partial R-squared of excluded instruments:   0.0040
Test of excluded instruments:
  F(  2,  3744) =     6.78
  Prob > F      =   0.0012


Summary results for first-stage regressions
-------------------------------------------

Variable     | Shea Partial R2 |   Partial R2    |  F(  2,  3744)    P-value
link         |     0.0040      |     0.0040      |        6.78      0.0012

NB: first-stage F-stat heteroskedasticity-robust

Underidentification tests
Ho: matrix of reduced form coefficients has rank=K1-1 (underidentified)
Ha: matrix has rank=K1 (identified)
Kleibergen-Paap rk LM statistic          Chi-sq(2)=13.50    P-val=0.0012
Kleibergen-Paap rk Wald statistic        Chi-sq(2)=13.73    P-val=0.0010

Weak identification test
Ho: equation is weakly identified
Kleibergen-Paap Wald rk F statistic                6.78
See main output for Cragg-Donald weak id test critical values
```

Figure A.3: Effect of Linking Contacts on Employment - Stata output page 2

```
Weak-instrument-robust inference
Tests of joint significance of endogenous regressors B1 in main equation
Ho: B1=0 and overidentifying restrictions are valid
Anderson-Rubin Wald test     F(2,3744)=3.67      P-val=0.0256
Anderson-Rubin Wald test     Chi-sq(2)=7.43      P-val=0.0244
Stock-Wright LM S statistic  Chi-sq(2)=7.21      P-val=0.0272

NB: Underidentification, weak identification and weak-identification-robust
    test statistics heteroskedasticity-robust

Number of observations              N  =       3791
Number of regressors                K  =         46
Number of instruments               L  =         47
Number of excluded instruments      L1 =          2


2-Step GMM estimation
---------------------

Estimates efficient for arbitrary heteroskedasticity
Statistics robust to heteroskedasticity

                                                      Number of obs =     3791
                                                      F( 45,  3745) =    15.89
                                                      Prob > F      =   0.0000
Total (centered) SS     =  918078.3596                Centered R2   =  -0.0091
Total (uncentered) SS   =  6895810.873                Uncentered R2 =   0.8657
Residual SS             =  926418.2005                Root MSE      =    15.63

------------------------------------------------------------------------------
             |               Robust
    hrswork2 |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
        link |   18.86574   8.249397     2.29   0.022     2.697217    35.03426
        ager |   .8656213   .2261097     3.83   0.000     .4224544    1.308788
       agesq |  -.0098397    .002543    -3.87   0.000    -.0148238   -.0048556
        male |   4.023942   .8698108     4.63   0.000     2.319144     5.72874
     hschool |  -.5159313   1.067272    -0.48   0.629    -2.607746    1.575883
     college |   -1.50408   1.618448    -0.93   0.353    -4.676179     1.66802
        univ |   1.086447   2.395214     0.45   0.650    -3.608086    5.780981
       kids2 |  -.8363128   .9127021    -0.92   0.360    -2.625176    .9525504
       kids3 |  -1.585415    .948024    -1.67   0.094    -3.443508    .2726784
       kids4 |  -2.993149   1.327609    -2.25   0.024    -5.595216   -.3910826
       kids5 |  -2.177715    1.78965    -1.22   0.224    -5.685365    1.329936
  marstatus1 |  -.0001165   .8413107    -0.00   1.000    -1.649055    1.648822
  marstatus2 |  -.8436164   .9488623    -0.89   0.374    -2.703352     1.01612
     treat11 |  -3.356772   4.489924    -0.75   0.455    -12.15686    5.443317
        occ2 |  -3.717469   1.418402    -2.62   0.009    -6.497486   -.9374522
        occ3 |   1.554851   1.991001     0.78   0.435    -2.347438    5.457141
        occ4 |  -.3512335   1.616206    -0.22   0.828    -3.518938    2.816471
        occ5 |  -5.586473   1.432385    -3.90   0.000    -8.393896    -2.77905
        occ6 |  -6.376619    2.52118    -2.53   0.011    -11.31804   -1.435196
        occ7 |  -6.380069    1.37948    -4.62   0.000    -9.083801   -3.676337
        occ8 |    .921595   1.573621     0.59   0.558    -2.162646    4.005836
        occ9 |   7.745776   2.987816     2.59   0.010     1.889764    13.60179
       occ10 |   3.568073   2.507298     1.42   0.155    -1.346141    8.482288
        ind2 |   8.400747   3.565859     2.36   0.018     1.411791     15.3897
        ind3 |  -.5193776   1.709633    -0.30   0.761    -3.870196    2.831441
        ind4 |   .7939084   2.014428     0.39   0.693    -3.154298    4.742115
        ind5 |   1.612197   1.633866     0.99   0.324    -1.590121    4.814515
```

Figure A.4: Effect of Linking Contacts on Employment - Stata output page 3

```
        ind6 |   .8144642   1.304416     0.62   0.532    -1.742145    3.371073
        ind7 |  -1.422067   1.611749    -0.88   0.378    -4.581036    1.736903
        ind8 |   1.364481   1.620866     0.84   0.400    -1.812358    4.541319
        ind9 |   .7787832    1.87511     0.42   0.678    -2.896365    4.453932
     pension |  -2.520045   1.346258    -1.87   0.061    -5.158663     .118572
     welfare |   1.194978   .6789542     1.76   0.078    -.1357479    2.525704
       town2 |   1.300823   1.348244     0.96   0.335    -1.341686    3.943332
       town3 |  -2.233564   1.791849    -1.25   0.213    -5.745523    1.278395
       town4 |   .2738637   4.773839     0.06   0.954    -9.082689    9.630416
       town5 |  -.2999853   4.924362    -0.06   0.951    -9.951558    9.351588
       town6 |  -2.390795   4.795325    -0.50   0.618    -11.78946    7.007869
       town7 |   1.000292   1.507082     0.66   0.507    -1.953535    3.954119
       town8 |  -2.883684   4.466567    -0.65   0.519    -11.63799    5.870626
       town9 |  -.9954055   4.714996    -0.21   0.833    -10.23663    8.245817
      town10 |   .6152968   1.357843     0.45   0.650    -2.046026     3.27662
      town11 |   -.497202   1.573825    -0.32   0.752    -3.581842    2.587438
      town12 |  -3.957221   4.826424    -0.82   0.412    -13.41684    5.502396
      town13 |   3.114712   4.919707     0.63   0.527    -6.527737    12.75716
       _cons |   18.33207   5.733168     3.20   0.001     7.095267    29.56887
------------------------------------------------------------------------------
Underidentification test (Kleibergen-Paap rk LM statistic):             13.502
                                                   Chi-sq(2) P-val =    0.0012
------------------------------------------------------------------------------


Weak identification test (Kleibergen-Paap rk Wald F statistic):          6.778
Stock-Yogo weak ID test critical values: 10% maximal IV size             19.93
                                         15% maximal IV size             11.59
                                         20% maximal IV size              8.75
                                         25% maximal IV size              7.25
Source: Stock-Yogo (2005).  Reproduced by permission.
NB: Critical values are for Cragg-Donald F statistic and i.i.d. errors.
------------------------------------------------------------------------------
Hansen J statistic (overidentification test of all instruments):         0.830
                                                   Chi-sq(1) P-val =    0.3624
-orthog- option:
Hansen J statistic (eqn. excluding suspect orthog. conditions):          0.000
                                                   Chi-sq(0) P-val =        .
C statistic (exogeneity/orthogonality of suspect instruments):           0.830
                                                   Chi-sq(1) P-val =    0.3624
Instruments tested:   treat2
-endog- option:
Endogeneity test of endogenous regressors:                               5.073
                                                   Chi-sq(1) P-val =    0.0243
Regressors tested:    link
------------------------------------------------------------------------------
Instrumented:         link
Included instruments: ager agesq male hschool college univ kids2 kids3 kids4
                      kids5 marstatus1 marstatus2 treat11 occ2 occ3 occ4 occ5
                      occ6 occ7 occ8 occ9 occ10 ind2 ind3 ind4 ind5 ind6 ind7
                      ind8 ind9 pension welfare town2 town3 town4 town5 town6
                      town7 town8 town9 town10 town11 town12 town13
Excluded instruments: treat2 community5
Dropped collinear:    ind10
------------------------------------------------------------------------------
```

Figure A.5: Effect of Linking Contacts on Employment - Stata output page 4

Appendix B

Appendix of Tables and Figures for Chapter 3

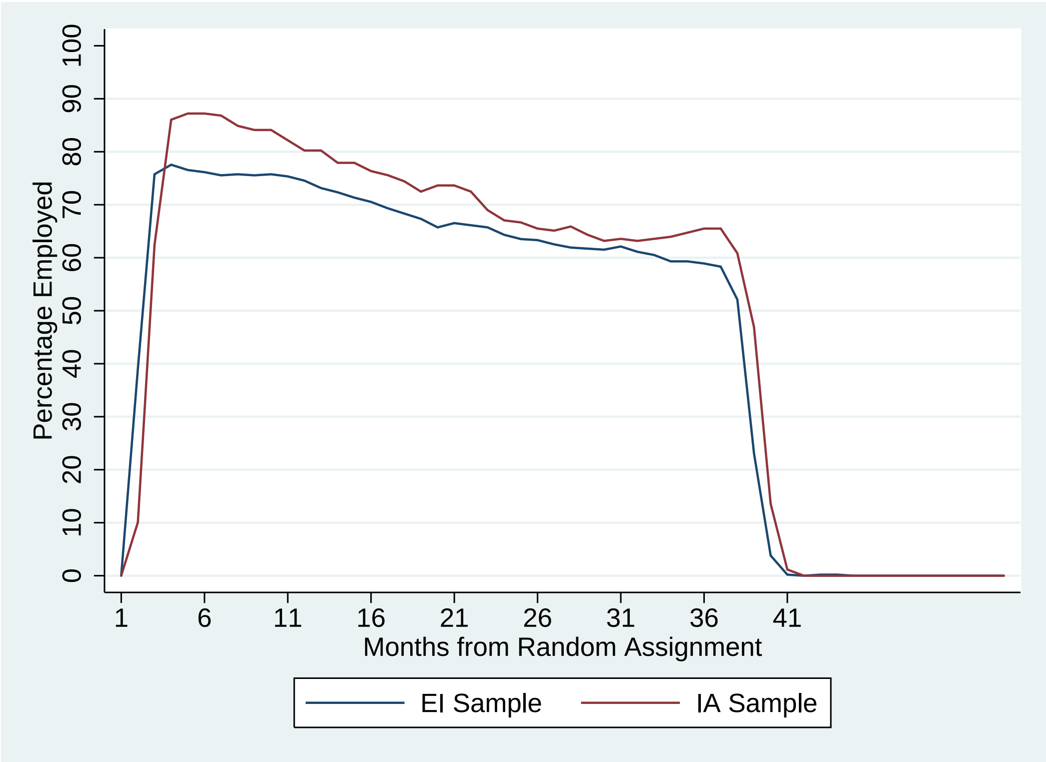


Figure B.1: Percentage of Program Group Members Actively Participating in CEIP, by Months from Enrolment

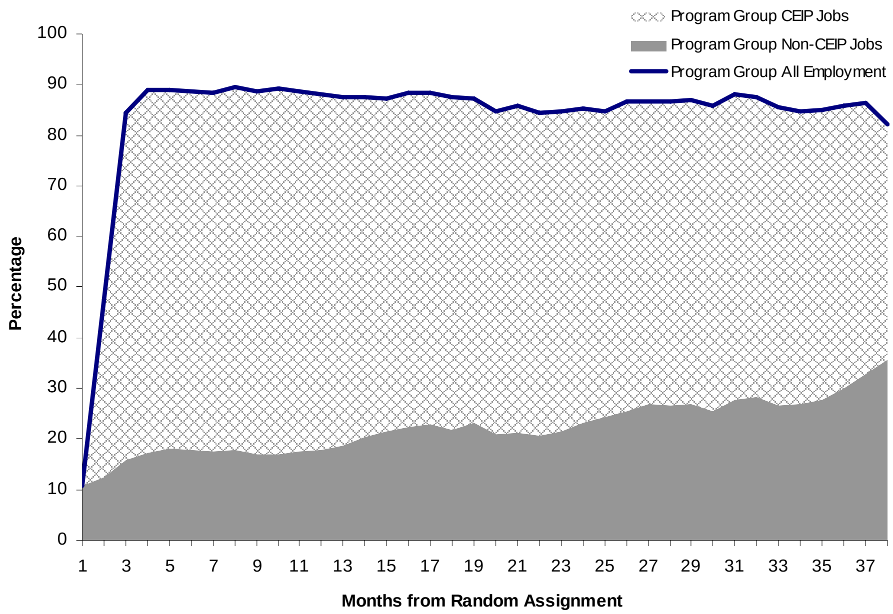


Figure B.2: CEIP vs. Non-CEIP Employment - EI Sample

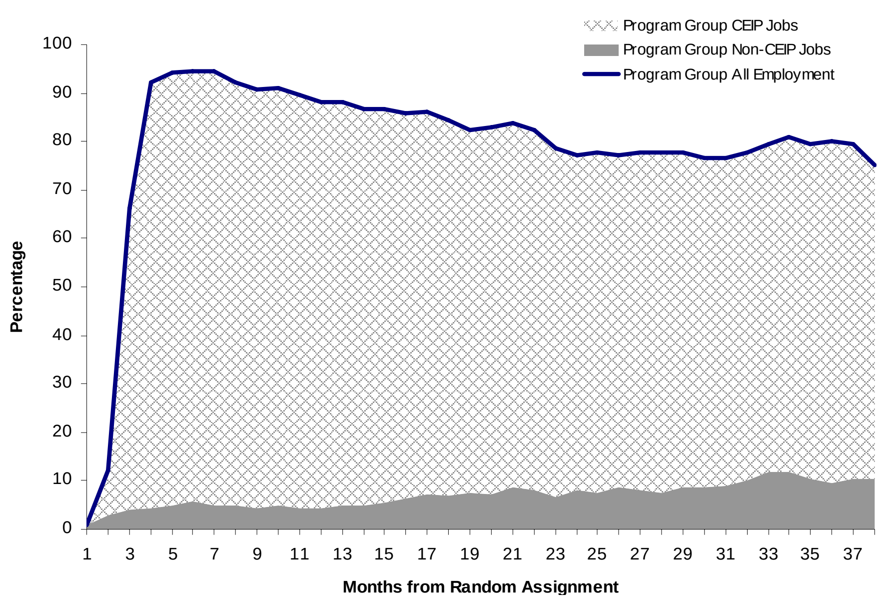


Figure B.3: CEIP vs. Non-CEIP Employment - EI Sample

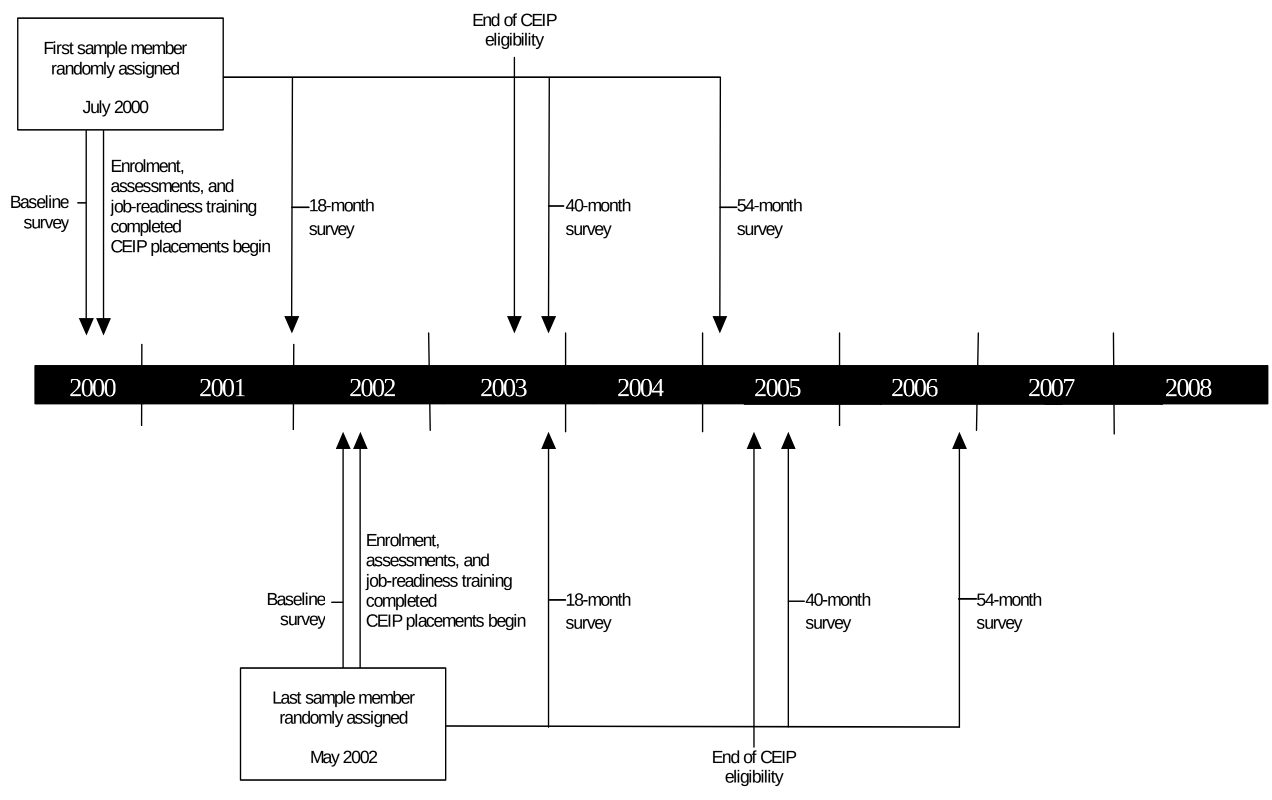


Figure B.4: Timeline for CEIP Participant Surveys

Table B.1: Summary Statistics

Variable	Mean	St. Dev.	Min.	Max.	Mean Control	Mean Program
Age	38.76	10.56	19	65	38.69	38.84
Male	0.5158	0.499	0	1	0.536	0.495
High school	0.507	0.499	0	1	0.504	0.510
College	0.1306	0.336	0	1	0.114	0.147
University	0.0358	0.185	0	1	0.042	0.029
Single	0.359	0.479	0	1	0.351	0.366
Married	0.446	0.497	0	1	0.441	0.450
Separated	0.087	0.282	0	1	0.098	0.077
Divorced/widowed	0.105	0.307	0	1	0.107	0.103
Years at current location	12.66	13.35	0.5	62	12.33	12.98
Relatives in Cape Breton	0.971	0.168	0	1	0.957	0.984
Household size	3.0	1.21	1	6	2.97	3.03
No. of firms worked for	1.633	0.967	0	5	1.62	1.64
Household income	27,106	17,953	1	140,000	26,564	27,630
job	0.438	0.496	0	1	0.394	0.481

Table B.2: Variable Definitions

Variable	Definition
Dependent variables	
forvol	How often volunteered on behalf of group/organization in last year
Independent variables	
bage	Age at baseline (in years)
bagesq	Age-squared
male	Dummy, 1 if male
Education - highest level achieved	
hschool	Has high school diploma
college	Has college diploma
univ	Has a university degree
kid1	Dummy, 1 if no children (**reference group**)
kid2	Dummy, 1 if 1 child
kid3	Dummy, 1 if 2 children
kid4	Dummy, 1 if 3 or more children
ei	Receiving Employment Insurance at time of enrolment
bsingle	Single at baseline
bmarried	Married or common-law at baseline
bsepdiv	Separated or divorced at baseline
breswage	Expected wage at baseline
hhincome	Pretax household income for the year (at baseline)
currentloc	How long lived at current address (in years)
relativesincb	Relatives living in Cape Breton, besides household (dummy)
hsize(1-6)	Number of persons in household (**hsize1: reference group**)
firm(1-5)	Companies worked for in last 5 years (**firm1: reference group**)
Instrument	
job	Dummy, 1 if employed full-time at the time of survey

Table B.3: Frequency of Volunteering

	Wave 1		Wave 2		Wave 3		Wave 4	
	P	C	P	C	P	C	P	C
How often did formal volunteering in past 12 months (percent)								
Never	51.2	46.1	53.1	62.8	52.8	66.7	55.3	65.0
Less than once a month	16.0	18.4	11.9	9.6	12.7	7.7	12.7	7.8
Once a month	10.7	10.9	13.0	10.6	10.4	8.9	10.0	8.6
Once a week	9.8	9.0	11.6	7.3	11.3	9.2	9.0	9.1
Few times a week	9.7	12.3	8.0	8.0	9.4	6.2	10.9	7.3
Everyday	2.5	3.3	2.3	1.7	3.4	1.3	2.0	2.2

P: Program group; C: Control group

Table B.4: Volunteer and donor rates, population aged 15 and older, Canada, 2007

	Number of volunteers (thousands)	Volunteer rate (percent)	Number of donors (thousands)	Donor rate (percent)
Canada	12,478	46.1	22,841	84.4
Nova Scotia	431	55.3	675	86.6

Source: Hall et al. (2009)

Table B.5: Volunteer rate by personal and economic characteristics, Nova Scotia 2007

	Volunteer rate (percent)	Average Annual volunteer hours (hours)	Population distribution (percent)	Percentage of total volunteer hours (percent)
Total	55.3	183	100.0	100.0
Age				
15 to 24	64.7	132	16.0	13.5
25 to 34	53.4	164	14.5	12.5
35 to 44	58.3	189	17.4	19.0
45 to 54	58.9	161	19.5	18.4
55 to 64	52.1	180	15.7	14.6
65 and older	43.9	298	17.0	22.0
Sex				
Male	52.6	182	48.6	46.3
Female	57.8	183	51.4	53.7
Marital Status				
Married or common-law	57.7	176	60.8	61.5
Single, never married	57.1	181	25.3	26.0
Separated or divorced	45.6	192	7.7	6.7
Education				
Less than high school	45.4	129	19.2	11.2
Graduated from high school	44.5	188	15.5	13.0
Some postsecondary	46.8	179	9.3	7.8
Postsecondary diploma	57.8	183	34.7	36.5
University degree	72.7	205	21.2	31.5
Labour force status				
Employed	59.8	164	62.7	60.9
Not in the labour force	51.3	211	34.8	37.5
Household income				
Less than $20,000	38.6	188	15.4	11.1
$20,000 to $39,999	43.4	168	21.9	15.8
$40,000 to $59,999	51.8	208	18.5	19.8
$60,000 to $79,999	65.8	206	16.4	22.0
$80,000 to $99,999	66.0	182	10.4	12.4
$100,000 or more	72.7	151	17.4	19.0
Presence of children in household				
No children	49.3	203	67.5	67.0
Pre-schooled aged children only	47.4	93	5.5	2.4
Pre-school & school aged children	63.5	195	3.6	-
School aged children only	73.3	154	23.4	26.2

Source: Hall et al. (2009)

Table B.6: Impact of Employment on Formal Volunteering: Probit in First Stage

VARIABLES	Wave 2 CF	Wave 2	Wave 3 CF	Wave 3	Wave 4 CF	Wave 4
job	**0.736*** **	**0.0475**	**-5.173*** **	**-0.170** **	**-4.817** **	**-0.0805**
	(0.237)	**(0.0744)**	**(1.140)**	**(0.0867)**	**(2.410)**	**(0.0958)**
bage	-0.0136	-0.00371	0.101**	-0.0423	0.173	-0.0198
	(0.0296)	(0.0290)	(0.0447)	(0.0289)	(0.105)	(0.0323)
bagesq	0.000271	0.000129	-0.00132**	0.000738**	-0.00239	0.000388
	(0.000368)	(0.000358)	(0.000600)	(0.000355)	(0.00150)	(0.000407)
male	-0.208***	-0.201***	-0.241***	-0.325***	-0.251**	-0.284***
	(0.0775)	(0.0772)	(0.0836)	(0.0816)	(0.108)	(0.106)
hschool	0.300***	0.273***	0.238**	0.154*	0.123	0.265**
	(0.0844)	(0.0834)	(0.0952)	(0.0934)	(0.136)	(0.115)
college	0.225**	0.227**	0.583***	0.196	0.511**	0.256*
	(0.108)	(0.107)	(0.169)	(0.140)	(0.206)	(0.145)
univ	0.578***	0.517***	0.905***	0.353**	0.755***	0.509**
	(0.159)	(0.161)	(0.203)	(0.178)	(0.252)	(0.202)
kid2	0.102	0.0768	0.121	0.0808	0.0410	0.189*
	(0.104)	(0.105)	(0.108)	(0.106)	(0.131)	(0.111)
kid3	0.209*	0.222*	-0.0268	0.299**	-0.163	0.323**
	(0.119)	(0.119)	(0.146)	(0.124)	(0.280)	(0.145)
kid4	0.354**	0.363**	0.289	0.535***	0.117	0.434*
	(0.167)	(0.169)	(0.197)	(0.189)	(0.270)	(0.236)
breswage	0.0215**	0.0173*	0.0240	0.00697	0.0251*	0.00954
	(0.00951)	(0.00952)	(0.0151)	(0.0148)	(0.0149)	(0.0135)
bmarried	0.176	0.194*	0.149	-0.0806	0.197	0.133
	(0.113)	(0.113)	(0.138)	(0.119)	(0.150)	(0.141)
bsepdiv	0.0915	0.106	0.251*	-0.124	-0.189	-0.146
	(0.114)	(0.114)	(0.151)	(0.122)	(0.138)	(0.141)
hhincome	-1.33e-05*	-3.81e-06	5.90e-05***	1.20e-05**	5.13e-05**	9.46e-06
	(7.41e-06)	(6.82e-06)	(1.22e-05)	(5.92e-06)	(2.27e-05)	(7.25e-06)
hhincomesq	1.44e-10*	5.63e-11	-3.73e-10***	-7.93e-11	-3.84e-10**	-7.05e-11
	(7.84e-11)	(7.43e-11)	(8.34e-11)	(5.59e-11)	(1.80e-10)	(7.69e-11)
Observations	1162	1162	1055	1055	875	875
ll	-1514	-1519	-1325	-1336	-1117	-1120
Pseudo R^2	0.0242	0.021	0.037	0.028	0.0229	0.0208

Standard errors, in parentheses, bootstrapped for all columns. *** p<0.01, ** p<0.05, * p<0.1

Table B.7: Impact of Employment on Formal Volunteering: LPM in First Stage

VARIABLES	Wave 2 CF	Wave 2	Wave 3 CF	Wave 3	Wave 4 CF	Wave 4
job	**0.658*** **	**0.0475**	**-4.083*** **	**-0.170** **	**-5.642*** **	**-0.0805**
	(0.190)	**(0.0744)**	**(0.885)**	**(0.0867)**	**(2.073)**	**(0.0958)**
bage	-0.0151	-0.00371	0.105**	-0.0423	0.274**	-0.0198
	(0.0297)	(0.0290)	(0.0446)	(0.0289)	(0.112)	(0.0323)
bagesq	0.000292	0.000129	-0.00137**	0.000738**	-0.00387**	0.000388
	(0.000370)	(0.000358)	(0.000597)	(0.000355)	(0.00160)	(0.000407)
male	-0.207***	-0.201***	-0.238***	-0.325***	-0.231**	-0.284***
	(0.0772)	(0.0772)	(0.0834)	(0.0816)	(0.109)	(0.106)
hschool	0.306***	0.273***	0.237**	0.154*	0.0377	0.265**
	(0.0844)	(0.0834)	(0.0948)	(0.0934)	(0.142)	(0.115)
college	0.224**	0.227**	0.605***	0.196	0.655***	0.256*
	(0.108)	(0.107)	(0.169)	(0.140)	(0.225)	(0.145)
univ	0.588***	0.517***	0.925***	0.353**	0.887***	0.509**
	(0.158)	(0.161)	(0.203)	(0.178)	(0.267)	(0.202)
kid2	0.108	0.0768	0.123	0.0808	-0.0407	0.189*
	(0.104)	(0.105)	(0.108)	(0.106)	(0.133)	(0.111)
kid3	0.210*	0.222*	-0.0367	0.299**	-0.434	0.323**
	(0.119)	(0.119)	(0.147)	(0.124)	(0.301)	(0.145)
kid4	0.355**	0.363**	0.277	0.535***	-0.0394	0.434*
	(0.167)	(0.169)	(0.197)	(0.189)	(0.278)	(0.236)
breswage	0.0219**	0.0173*	0.0240	0.00697	0.0333**	0.00954
	(0.00944)	(0.00952)	(0.0150)	(0.0148)	(0.0153)	(0.0135)
bmarried	0.173	0.194*	0.157	-0.0806	0.236	0.133
	(0.113)	(0.113)	(0.138)	(0.119)	(0.151)	(0.141)
bsepdiv	0.0898	0.106	0.258*	-0.124	-0.222	-0.146
	(0.114)	(0.114)	(0.149)	(0.122)	(0.140)	(0.141)
hhincome	-1.49e-05**	-3.81e-06	6.13e-05***	1.20e-05**	7.43e-05***	9.46e-06
	(7.59e-06)	(6.82e-06)	(1.25e-05)	(5.92e-06)	(2.50e-05)	(7.25e-06)
hhincomesq	1.58e-10**	5.63e-11	-3.89e-10***	-7.93e-11	-5.57e-10***	-7.05e-11
	(7.97e-11)	(7.43e-11)	(8.52e-11)	(5.59e-11)	(1.96e-10)	(7.69e-11)
Observations	1162	1162	1055	1055	875	875
ll	-1513	-1519	-1325	-1336	-1115	-1120
R^2	0.025	0.021	0.037	0.028	0.0246	0.0208

Standard errors, in parentheses, bootstrapped for all columns. *** p<0.01, ** p<0.05, * p<0.1

Table B.8: Estimation of Probability of Employment: First-Stage Probit Results

VARIABLES	Wave 1	Wave 2	Wave 3	Wave 4
treat	**0.0477**	**1.170*** **	**-0.239*** **	**-0.119**
	(0.0832)	**(0.0898)**	**(0.0817)**	**(0.0904)**
bage	0.0285	0.0763**	0.108***	0.157***
	(0.0310)	(0.0307)	(0.0308)	(0.0344)
bagesq	-0.000379	-0.00105***	-0.00156***	-0.00228***
	(0.000395)	(0.000376)	(0.000387)	(0.000429)
male	-0.598***	0.0561	0.0799	0.0322
	(0.0936)	(0.0962)	(0.0930)	(0.102)
hschool	-0.123	-0.0891	0.0508	-0.119
	(0.0947)	(0.0955)	(0.0944)	(0.105)
college	0.0824	0.122	0.266*	0.189
	(0.133)	(0.150)	(0.137)	(0.148)
univ	-0.195	-0.405*	0.429*	0.207
	(0.219)	(0.244)	(0.227)	(0.250)
breswage	-0.0127	-0.0351***	0.0147	0.0130
	(0.0140)	(0.0135)	(0.0130)	(0.0157)
bmarried	0.0616	0.162	0.162	0.0415
	(0.129)	(0.130)	(0.126)	(0.138)
bsepdiv	-0.0106	0.131	0.293**	-0.0225
	(0.141)	(0.136)	(0.140)	(0.155)
hhincome	-1.43e-06	4.53e-05***	3.53e-05***	3.35e-05***
	(1.30e-05)	(8.34e-06)	(6.38e-06)	(7.31e-06)
hhincomesq	-1.13e-10	-4.09e-10***	-2.22e-10***	-2.52e-10***
	(2.35e-10)	(9.75e-11)	(6.08e-11)	(7.26e-11)
Constant	-2.212***	-1.772**	-2.819***	-3.471***
	(0.749)	(0.707)	(0.699)	(0.768)
Observations	1386	1165	1059	878
ll	-593.3	-577.9	-662.7	-540.9
Pseudo R^2	0.079	0.2177	0.089	0.1125

Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table B.9: Estimation of Probability of Employment: First-Stage OLS Results

VARIABLES	Wave 1	Wave 2	Wave 3	Wave 4
treat	**0.00989**	**0.366*** **	**-0.0872*** **	**-0.0435**
	(0.0204)	**(0.0250)**	**(0.0296)**	**(0.0326)**
bage	0.00721	0.0228**	0.0368***	0.0523***
	(0.00702)	(0.00922)	(0.0102)	(0.0114)
bagesq	-9.32e-05	-0.000313***	-0.000529***	-0.000758***
	(8.64e-05)	(0.000113)	(0.000125)	(0.000137)
male	-0.146***	0.00869	0.0287	0.0119
	(0.0233)	(0.0270)	(0.0338)	(0.0367)
hschool	-0.0317	-0.0301	0.0171	-0.0443
	(0.0224)	(0.0274)	(0.0341)	(0.0375)
college	0.0235	0.0395	0.0987**	0.0674
	(0.0377)	(0.0416)	(0.0502)	(0.0522)
univ	-0.0442	-0.120*	0.154*	0.0727
	(0.0561)	(0.0725)	(0.0802)	(0.0861)
breswage	-0.00231	-0.00990**	0.00496	0.00451
	(0.00246)	(0.00453)	(0.00439)	(0.00568)
bmarried	0.0195	0.0442	0.0575	0.0155
	(0.0320)	(0.0370)	(0.0459)	(0.0484)
bsepdiv	0.00485	0.0351	0.0993**	-0.0137
	(0.0349)	(0.0391)	(0.0502)	(0.0547)
hhincome	-7.01e-07	1.42e-05***	1.29e-05***	1.18e-05***
	(3.07e-06)	(2.50e-06)	(2.29e-06)	(2.47e-06)
hhincomesq	-0	-1.26e-10***	-8.22e-11***	-8.88e-11***
	(5.21e-11)	(0)	(0)	(0)
Constant	-0.0637	-0.0546	-0.471**	-0.656**
	(0.160)	(0.210)	(0.239)	(0.258)
Observations	1386	1165	1059	878
R^2	0.068	0.257	0.121	0.143
rmse	0.371	0.411	0.474	0.471
ll	-580.0	-601.9	-697.0	-569.6

Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table B.10: Impact of Employment on Informal Volunteering: Probit in First Stage

VARIABLES	Wave 2 CF	Wave 2	Wave 3 CF	Wave 3	Wave 4 CF	Wave 4
job	**0.196**	**-0.142****	**-0.867**	**-0.148****	**-1.941**	**-0.0745**
	(0.213)	**(0.0716)**	**(0.900)**	**(0.0671)**	**(1.850)**	**(0.0785)**
bage	-0.00700	-0.00189	0.0306	0.00970	0.0771	0.000635
	(0.0247)	(0.0243)	(0.0360)	(0.0243)	(0.0840)	(0.0275)
bagesq	6.80e-05	-4.74e-06	-0.000468	-0.000169	-0.00124	-0.000139
	(0.000304)	(0.000298)	(0.000484)	(0.000302)	(0.00118)	(0.000334)
male	0.191***	0.195***	0.0990	0.0856	0.0547	0.0419
	(0.0716)	(0.0722)	(0.0788)	(0.0759)	(0.0868)	(0.0870)
hschool	0.183**	0.171**	0.122	0.111	0.00658	0.0658
	(0.0740)	(0.0735)	(0.0856)	(0.0849)	(0.110)	(0.0964)
college	0.244**	0.244**	0.0736	0.0182	0.0285	-0.0708
	(0.108)	(0.107)	(0.147)	(0.117)	(0.157)	(0.129)
univ	-0.0142	-0.0427	0.273	0.194	0.0728	-0.0218
	(0.201)	(0.201)	(0.243)	(0.222)	(0.204)	(0.184)
kid2	0.169*	0.157	-0.00501	-0.00916	0.0387	0.0981
	(0.0976)	(0.0976)	(0.102)	(0.102)	(0.110)	(0.0998)
kid3	0.222**	0.228**	0.0182	0.0666	-0.0363	0.156
	(0.112)	(0.112)	(0.140)	(0.123)	(0.231)	(0.141)
kid4	0.412**	0.415**	0.00477	0.0436	-0.00216	0.123
	(0.186)	(0.187)	(0.188)	(0.181)	(0.245)	(0.205)
breswage	-0.00338	-0.00525	0.00965	0.00726	0.0186	0.0126
	(0.0111)	(0.0111)	(0.0137)	(0.0132)	(0.0121)	(0.0104)
bmarried	0.0202	0.0299	0.0279	-0.00457	0.0360	0.0123
	(0.0948)	(0.0953)	(0.111)	(0.102)	(0.114)	(0.114)
bsepdiv	0.0587	0.0669	0.117	0.0631	0.176	0.193
	(0.116)	(0.116)	(0.136)	(0.120)	(0.150)	(0.148)
hhincome	1.49e-06	6.19e-06	1.09e-05	4.11e-06	1.09e-05	-5.53e-06
	(7.10e-06)	(6.25e-06)	(1.08e-05)	(5.88e-06)	(1.76e-05)	(6.62e-06)
hhincomesq	-0	-6.72e-11	-6.43e-11	-0	-1.11e-10	0
	(7.46e-11)	(6.73e-11)	(7.77e-11)	(5.14e-11)	(1.39e-10)	(6.44e-11)
Observations	1160	1160	1045	1045	869	869
ll	-1938	-1939	-1687	-1687	-1421	-1422
Pseudo R^2	0.0108	0.0099	0.0087	0.0086	0.0115	0.0111

Standard errors, in parentheses, bootstrapped for all columns. *** p<0.01, ** p<0.05, * p<0.1

Table B.11: Reasons for Volunteering at Wave 2 (18-month) Survey

Reasons for Volunteering	% of participants
To help cause in which personally believe	15.7
Because friends volunteer	3.2
To improve job skills	1.2
To improve job opportunities	1.7
To fulfill religious obligations/beliefs	2.7
Enjoy helping other people	85.9
Required by school/employer/government	0.8
Already work for volunteer organization	1.4
Something to do	4.8
No of observations	1062

Table B.12: Forms of Volunteering (in percentages)

Formal Volunteering	Wave 2	Wave 3	Wave 4
Did canvassing, campaigning or fundraising	54.4	46.7	56.7
Serve as an unpaid member of a Board or Committee	32.7	28.7	35.1
Provide information/help to educate/influence public opinion	29.7	28.1	32.5
Organize or supervise activities for an organization	60.4	58.4	59.2
Did consulting, executive, office or admin work	27.9	24.4	27.5
Teach or coach for an organization	27.7	25.1	24.0
Provide care or support, inc counseling & friendly visits	27.9	27.3	31.4
Collect/serve/deliver food as a volunteer through an orgn	36.6	31.1	38.0
Did volunteer driving on behalf of an organization	27.4	19.2	25.1
Did formal volunteering through some other way	35.9	25.7	38.0
Observations	566	505	458

B.0.2 STATA Output

We include here complete output from STATA for the impact of employment on levels of formal volunteering, using probit estimation in the first stage. A summary of these results have been presented in columns Wave 2a, Wave 3a and Wave 4a, respectively, of Table B.6.

```
oprobit forvol jobmon jobres2 `control' if wave==2, nolog
 vce(bootstrap, reps(200) seed(123456))
(running oprobit on estimation sample)

Ordered probit regression                         Number of obs      =      1162
                                                  Replications       =       200
                                                  Wald chi2(27)      =         .
                                                  Prob > chi2        =         .
Log likelihood = -1514.3353                       Pseudo R2          =    0.0252

------------------------------------------------------------------------------
             |   Observed   Bootstrap                         Normal-based
      forvol |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
         job |   .7356077   .2367827     3.11   0.002     .2715222    1.199693
     jobres2 |  -.3272099   .1100562    -2.97   0.003     -.542916   -.1115037
        bage |  -.0136307   .0295964    -0.46   0.645    -.0716386    .0443773
      bagesq |   .0002711   .0003683     0.74   0.462    -.0004507    .0009929
        male |  -.2082515   .0774923    -2.69   0.007    -.3601336   -.0563694
     hschool |   .2997169   .0844437     3.55   0.000     .1342102    .4652235
     college |   .2246232    .107607     2.09   0.037     .0137172    .4355291
        univ |   .5784839   .1590217     3.64   0.000     .2668072    .8901606
        kid2 |   .1022922   .1042013     0.98   0.326    -.1019386    .3065231
        kid3 |   .2087189    .119107     1.75   0.080    -.0247266    .4421644
        kid4 |   .3535474   .1673821     2.11   0.035     .0254846    .6816102
    breswage |   .0215272   .0095099     2.26   0.024     .0028882    .0401662
    bmarried |   .1757424    .113147     1.55   0.120    -.0460217    .3975065
     bsepdiv |   .0915196   .1138864     0.80   0.422    -.1316938    .3147329
          ei |   .1865051   .1086589     1.72   0.086    -.0264623    .3994726
    hhincome |  -.0000133   7.41e-06    -1.80   0.073    -.0000278    1.22e-06
  hhincomesq |   1.44e-10   7.84e-11     1.83   0.067    -9.92e-12    2.97e-10
yearsatcur~c |   .0036433   .0032874     1.11   0.268    -.0027998    .0100864
relativesi~b |    .256804   .3526454     0.73   0.466    -.4343684    .9479764
      hsize2 |  -.2540323   .1600708    -1.59   0.113    -.5677654    .0597008
      hsize3 |   -.128722   .1567402    -0.82   0.412    -.4359272    .1784831
      hsize4 |  -.2397727   .1690585    -1.42   0.156    -.5711212    .0915757
      hsize5 |  -.1607374   .2201278    -0.73   0.465    -.5921799    .2707051
      hsize6 |  -.5566098   .2842984    -1.96   0.050    -1.113824    .0006048
       firm2 |  -.0471813   .1569885    -0.30   0.764     -.354873    .2605105
       firm3 |  -.0906727   .1595198    -0.57   0.570    -.4033258    .2219805
       firm4 |   .0139947   .1888561     0.07   0.941    -.3561564    .3841458
       firm5 |   .2368543   .2664901     0.89   0.374    -.2854567    .7591654
-------------+----------------------------------------------------------------
       /cut1 |   1.019082   .7100676                     -.3726252    2.410789
       /cut2 |   1.320942   .7112234                     -.0730299    2.714915
       /cut3 |   1.721023   .7094537                      .3305194    3.111527
       /cut4 |   2.155235   .7137785                       .756255    3.554215
       /cut5 |   2.932799    .722427                      1.516868     4.34873
------------------------------------------------------------------------------
```

Figure B.5: Impact of Employment on Formal Volunteering: Stata output for Wave-2

```
oprobit forvol jobmon jobres3 `control' if wave==3, nolog
 vce(bootstrap, reps(200) seed(123456))
(running oprobit on estimation sample)

Ordered probit regression                       Number of obs      =      1055
                                                Replications       =       199
                                                Wald chi2(27)      =         .
                                                Prob > chi2        =         .
Log likelihood = -1325.0852                     Pseudo R2          =    0.0386

------------------------------------------------------------------------------
             |   Observed   Bootstrap                         Normal-based
      forvol |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
         job |  -5.172943   1.139687    -4.54   0.000    -7.406689   -2.939197
     jobres3 |   2.108469   .4798174     4.39   0.000     1.168044    3.048894
        bage |   .1012619   .0447446     2.26   0.024      .013564    .1889597
      bagesq |  -.0013233   .0006003    -2.20   0.028    -.0024999   -.0001467
        male |  -.2411549   .0836151    -2.88   0.004    -.4050375   -.0772723
     hschool |   .2377832   .0951539     2.50   0.012      .051285    .4242814
     college |   .5831608   .1686715     3.46   0.001     .2525706    .9137509
        univ |   .9051562   .2025598     4.47   0.000     .5081464    1.302166
        kid2 |   .1206618   .1076182     1.12   0.262     -.090266    .3315895
        kid3 |  -.0267814    .146141    -0.18   0.855    -.3132125    .2596496
        kid4 |   .2894733    .196869     1.47   0.141    -.0963828    .6753295
    breswage |    .023969   .0150741     1.59   0.112    -.0055758    .0535138
    bmarried |   .1494257   .1375673     1.09   0.277    -.1202012    .4190526
     bsepdiv |   .2509752   .1512717     1.66   0.097     -.045512    .5474624
          ei |   .1465916    .103917     1.41   0.158     -.057082    .3502653
    hhincome |    .000059   .0000122     4.85   0.000     .0000351    .0000828
  hhincomesq |  -3.73e-10   8.34e-11    -4.47   0.000    -5.36e-10   -2.09e-10
yearsatcur~c |  -.0047913   .0034648    -1.38   0.167    -.0115822    .0019995
relativesi~b |   .2020823   .2951846     0.68   0.494    -.3764689    .7806335
      hsize2 |  -.3006409   .1718571    -1.75   0.080    -.6374747    .0361929
      hsize3 |   .0270674   .1695016     0.16   0.873    -.3051497    .3592844
      hsize4 |   .0461285   .1986865     0.23   0.816    -.3432899    .4355468
      hsize5 |  -.1332249   .2296767    -0.58   0.562    -.5833831    .3169332
      hsize6 |   .1619143     .26485     0.61   0.541    -.3571822    .6810108
       firm2 |  -.1427149   .1552637    -0.92   0.358    -.4470262    .1615964
       firm3 |   .1234003   .1901686     0.65   0.516    -.2493233     .496124
       firm4 |   .4067396   .2386848     1.70   0.088    -.0610741    .8745533
       firm5 |    .400852   .2719553     1.47   0.140    -.1321706    .9338745
-------------+----------------------------------------------------------------
       /cut1 |   1.669299   .8003305                      .1006802    3.237918
       /cut2 |   1.955471   .8029314                      .3817547    3.529188
       /cut3 |   2.297678    .806654                      .7166649    3.878691
       /cut4 |   2.766423   .8055702                      1.187534    4.345312
       /cut5 |   3.499989   .8057639                      1.920721    5.079257
------------------------------------------------------------------------------
Note: one or more parameters could not be estimated in 1 bootstrap replicate;
      standard error estimates include only complete replications.
```

Figure B.6: Impact of Employment on Formal Volunteering: Stata output for Wave-3

```
oprobit forvol jobmon jobres4 `control' if wave==4, nolog
 vce(bootstrap, reps(200) seed(123456))
(running oprobit on estimation sample)

Ordered probit regression                         Number of obs      =       875
                                                  Replications       =       200
                                                  Wald chi2(27)      =         .
                                                  Prob > chi2        =         .
Log likelihood = -1116.9371                       Pseudo R2          =    0.0269

------------------------------------------------------------------------------
             |   Observed   Bootstrap                         Normal-based
      forvol |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
      job    |  -4.817234    2.41026    -2.00   0.046    -9.541257   -.0932114
     jobres4 |   1.989471    1.01454     1.96   0.050     .0010088    3.977933
        bage |   .1725462   .1051137     1.64   0.101    -.0334729    .3785654
      bagesq |   -.002394   .0014987    -1.60   0.110    -.0053314    .0005435
        male |  -.2510776   .1082238    -2.32   0.020    -.4631923   -.0389629
     hschool |   .1234368    .136364     0.91   0.365    -.1438317    .3907054
     college |   .5106372   .2055209     2.48   0.013     .1078235    .9134508
        univ |   .7548307   .2523339     2.99   0.003     .2602653    1.249396
        kid2 |   .0409923   .1307062     0.31   0.754    -.2151872    .2971717
        kid3 |  -.1627647   .2799917    -0.58   0.561    -.7115383    .3860088
        kid4 |   .1167451   .2703075     0.43   0.666    -.4130479     .646538
    breswage |   .0251477   .0148578     1.69   0.091     -.003973    .0542684
    bmarried |   .1968903   .1496906     1.32   0.188    -.0964978    .4902785
     bsepdiv |  -.1887099    .138011    -1.37   0.172    -.4592064    .0817866
          ei |   .0661023   .1496889     0.44   0.659    -.2272826    .3594872
    hhincome |   .0000513   .0000227     2.27   0.023     6.92e-06    .0000958
  hhincomesq |  -3.84e-10   1.80e-10    -2.14   0.032    -7.36e-10   -3.25e-11
yearsatcur~c |  -.0045165   .0053087    -0.85   0.395    -.0149214    .0058884
relativesi~b |   .0821487   .2793427     0.29   0.769     -.465353    .6296504
      hsize2 |  -.1540696   .2264215    -0.68   0.496    -.5978476    .2897085
      hsize3 |  -.1182281   .2480687    -0.48   0.634    -.6044338    .3679777
      hsize4 |  -.1380004   .2308511    -0.60   0.550    -.5904603    .3144594
      hsize5 |  -.4671536   .3531897    -1.32   0.186    -1.159393    .2250856
      hsize6 |  -.6333646    .723284    -0.88   0.381    -2.050975    .7842461
       firm2 |   .3859697   .2291387     1.68   0.092    -.0631338    .8350733
       firm3 |   .8712819   .3854021     2.26   0.024     .1159078    1.626656
       firm4 |   .8521854   .3448311     2.47   0.013     .1763289    1.528042
       firm5 |   .9562564   .4080433     2.34   0.019     .1565062    1.756007
-------------+----------------------------------------------------------------
       /cut1 |   2.582219   1.242628                      .1467123    5.017726
       /cut2 |   2.918473   1.243241                       .481766     5.35518
       /cut3 |   3.244633    1.24492                      .8046339    5.684632
       /cut4 |   3.638686   1.245142                      1.198251     6.07912
       /cut5 |   4.420196   1.257012                      1.956499    6.883894
------------------------------------------------------------------------------
```

Figure B.7: Impact of Employment on Formal Volunteering: Stata output for Wave-4

Appendix C

Appendix of Tables and Figures for Chapter 4

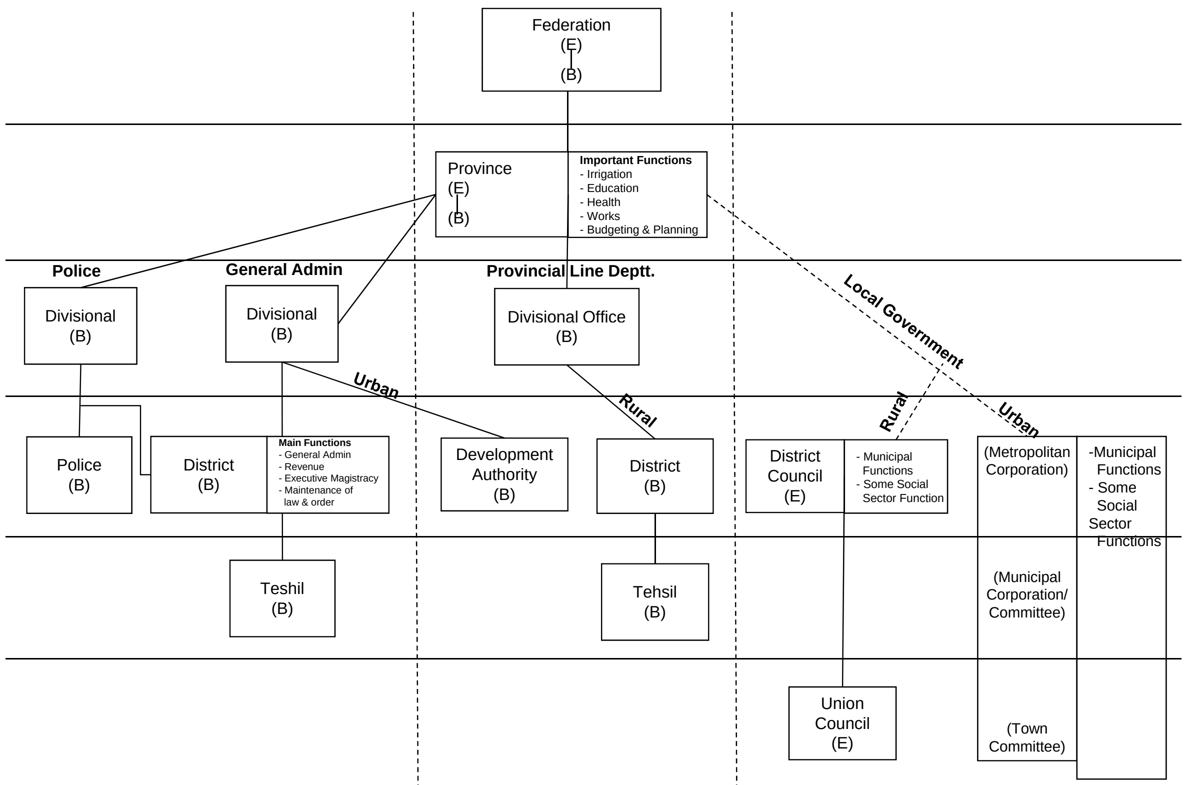


Figure C.1: Pre-decentralization Local Government Structure

E stands for Elected Politician, B stands for Bureaucrat.

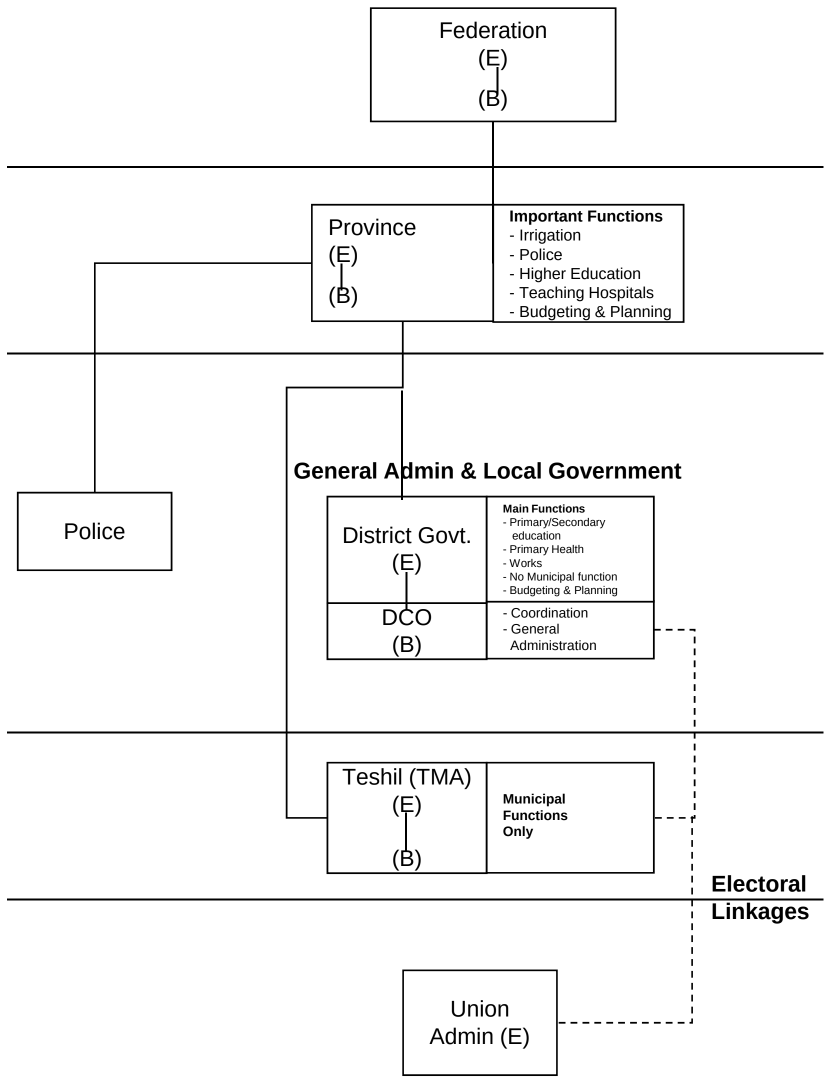


Figure C.2: Post-decentralization Local Government Structure
E stands for Elected Politician, B stands for Bureaucrat.

Table C.1: Classification of Decentralized and Non-Decentralized Sectors

Decentralized Sectors (District Governments)	**Non-decentralized Sectors** (Provincial Government)
Health	Health
Agriculture	Agriculture
Education	Education
Livestock	Livestock
	Police
	Irrigation
Fisheries	
Land Revenue	
Forests	
Excise and Taxation	
Works	
Public Health	
Miscellaneous	
General Administration	
Communication	
Forests	
Industries	
Rural Electrification	
Housing	

Decentralized sectors include some complete sectors, and others where only a subset of all the functions in the sector have been decentralized. The above classification covers the entire set of decentralized sectors, but only a small subset of non-decentralized sectors.

Table C.2: Summary Statistics of Aggregate District Expenditures

	2001-02	2002-03	2003-04
Establishment	29,231.71	36,580.74	44,536.2
	(3037.96)	(3747.81)	(4722.27)
	[1205]	[1259]	[1316]
Non-Establishment	6359.86	7609.91	6998.67
	(469.19)	(532.56)	(525.75)
	[1233]	[1279]	[1318]
Development	11,143.75	11,205.18	18,193.07
	(1012.91)	(1023.41)	(1437.07)
	[317]	[275]	[258]
Total Expenditures	46,735.32	55,395.83	69,724.94
	(3287.71)	(3980.11)	(5076.53)
	[2755]	[2813]	[2892]

Standard deviations in parentheses. Number of observations in square brackets.
Each observation represents expenditure on a specific sub-sector for a specific activity in a particular district for a given year.
All figures are nominal, in million rupees, and have been aggregated from district budgets.

Table C.3: Summary Statistics of Aggregate District Expenditures by Sectors - I

	Edu	Health	Agriculture	Electrif.	Livestock	Revenue
2001-02						
Establish	24,252.9	2,680.1	223.5	170.1	365.4	468.23
	(2734.9)	(290.9)	(32.3)	(22.4)	(29.7)	(60.96)
Non-Estab	1,858.1	2,684.6	112.1	135.5	29.75	41.51
	(186.4)	(342.4)	(23.6)	(53.9)	(2.1)	(20.2)
Development	2,108.37	532.92	190.27	1,069.1	3.83	0
	(321.5)	(107.6)	(79.4)	(270.2)	(3.83)	(0)
2002-03						
Establish	29,500.0	3,242.62	743.26	202.96	455.56	527.10
	(3375.5)	(360.7)	(65.0)	(27.1)	(52.7)	(70.1)
Non-Estab	1,915.95	2,795.05	264.11	195.42	65.06	61.32
	(199.5)	(363.7)	(37.5)	(81.8)	(8.8)	(32.8)
Development	1,271.80	668.19	331.16	825.0	14.53	0
	(244.8)	(295.7)	(97.3)	(170.3)	(9.8)	(0)
2003-04						
Establish	35,476.6	4,277.7	755.28	233.84	595.96	642.46
	(4285.1)	(536.8)	(67.2)	(28.2)	(55.8)	(86.5)
Non-Estab	1,476.93	2,316.57	247.09	25.71	100.88	24.26
	(2753.8)	(270.3)	(36.8)	(25.7)	(12.2)	(5.9)
Development	3,449.59	694.05	567.23	1,416.75	22.85	0
	(296.6)	(582.4)	(171.2)	(93.6)	(315.1)	(0)

Standard deviations in parentheses. All figures are nominal, in million rupees.
All figures have been aggregated from district budgets.

Table C.4: Summary Statistics of Aggregate District Expenditures by Sectors - II

	Fisheries	Indusry	Housing	Works	Adminis	Public Health
2001-02						
Establish	17.89	17.61	5.08	285.00	68.56	10.82
	(1.1)	(2.2)	(0.5)	(52.6)	(3.5)	(1.7)
Non-Estab	3.37	5.8	1.71	403.32	89.24	76.07
	(0.6)	(0.9)	(0.5)	(91.7)	(8.1)	(40.5)
Development	0.001	28.44	1.04	7,011.15	128.3	0
	(0)	(0)	(0.7)	(792.8)	(102.4)	(0)
2002-03						
Establish	18.2	19.55	6.07	304.87	777.68	12.54
	(1.1)	(2.8)	(1.1)	(34.1)	(88.1)	(2.2)
Non-Estab	9.70	7.46	0.97	492.64	419.38	107.18
	(2.3)	(1.1)	(0.3)	(141.6)	(52.1)	(56.2)
Development	1.5	16.93	86.23	6,796.49	833.50	0
	(1.5)	(0)	(0)	(795.7)	(231.9)	(0)
2003-04						
Establish	28.14	21.74	40.44	316.82	1,117.18	13.08
	(9.4)	(2.8)	(1.6)	(30.3)	(122.0)	(2.0)
Non-Estab	8.44	6.02	14.22	342.79	1,096.57	4.46
	(2.1)	(0.5)	(1.0)	(67.9)	(240.6)	(0.6)
Development	10.10	7.65	20.0	9,653.87	1,814.65	0
	(2.8)	(0)	(0)	(1030.1)	(366.1)	(0)

Standard deviations in parentheses. All figures are nominal, in million rupees.
All figures have been aggregated from district budgets.

Table C.5: Summary Statistics of Provincial Expenditures by Sectors

	Health	Agriculture	Livestock	Police
2001-02				
Establish	315.23	85.26	213.21	5,250.63
	(298.6)	(17.6)	(65.7)	(1329.1)
Non-Estab	3,031.6	22.98	129.73	1,274.83
	(1987)	(10.3)	(39.1)	(333.4)
Development	267.53	326.23	19.89	50.98
	(218.1)	(259.6)	(17.9)	(42.4)
2002-03				
Establish	370.29	96.36	178.64	5,658.54
	(344.6)	(20.3)	(53.2)	(1481.4)
Non-Estab	3,917.6	27.25	174.81	2,553.4
	(2503.7)	(13.2)	(67.9)	(1152.7)
Development	429.12	649.64	146.9	141.5
	(389.7)	(573.5)	(141.9)	(40.7)
2003-04				
Establish	513.78	110.51	204.57	6,901.40
	(438.5)	(24.5)	(59.7)	(1828.5)
Non-Estab	5,087.48	77.4	202.67	3,940.1
	(2717.6)	(62.1)	(81.3)	(2355.3)
Development	111.45	819.96	147.66	557.05
	(915.14)	(783.14)	(138.67)	(73.7)

Standard deviations in parentheses. All figures are nominal, in million rupees.
All figures have been aggregated from provincial budgets.

Table C.6: Variable Definitions

Variable	Definition
Dependent variable	
$lnexp$	log of expenditures
$reelection$	Dummy, 1 if nazim gets reelected in 2005 elections
Independent variables	
$post1$	Dummy, 1 for year-1 after decentralization
$post2$	Dummy, 1 for year-2 after decentralization
$treat1$	Dummy, 1 if expenditure type is development
$treat2$	Dummy, 1 if expenditure type is non-establishment
$post1 - treat1$	Dummy, 1 for change in development expenditure in year-1
$post2 - treat1$	Dummy, 1 for change in development expenditure in year-2
$post1 - treat2$	Dummy, 1 for change in non-establishment expenditure in year-1
$post2 - treat2$	Dummy, 1 for change in non-establishment expenditure in year-2
dev	Dummy, 1 if the sector (or functions within it) is decentralized
$physical$	Percentage increase in spending on physical infrastructure
$social$	Percentage increase in spending on Social sectors
$admin$	Percentage increase in spending on Administrative sectors
$householdsize$	Average household size in a district in 1998
$density$	Population density in the district in 1998
$paccahouses$	Proportion of *pacca* houses out of total houses in 1998
$literacy$	Literacy ratio 1998
$diversity$	Measure of ethnic diversity (fractionalization) in 1998
$urban$	Proportion of people living in urban areas in 1998

Table C.7: Difference-in-Difference Results

	(1) **DD-I**	(2) **DD-II**	(3) **DD-III**	(4) **DD-IV**
post1_treat1	**2.348*** **	**2.297*** **	**2.291*** **	**2.554*** **
(β_5)	(0.613)	(0.623)	(0.622)	(0.661)
post2_treat1	**5.373*** **	**5.314*** **	**5.284*** **	**5.526*** **
(β_7)	(0.481)	(0.485)	(0.491)	(0.518)
post1_treat2	**0.164*** **	**0.164*** **	**0.173*** **	**0.170*** **
(β_6)	(0.0323)	(0.0317)	(0.0332)	(0.035)
post2_treat2	**0.0963*** **	**0.0940** **	**0.114*** **	**0.117*** **
(β_8)	(0.0349)	(0.0350)	(0.0361)	(0.0381)
post1	0.0777***	-1.141***	-1.064***	-0.0710
(β_1)	(0.0192)	(0.0820)	(0.0850)	(0.209)
post2	0.194***	0.0459	0.0179	0.0166
(β_2)	(0.0299)	(0.0580)	(0.0601)	(0.219)
treat1	-3.283***	-3.283***	-3.721***	-3.671***
(β_3)	(0.511)	(0.513)	(0.549)	(0.544)
treat2	-1.244***	-1.244***	-1.272***	-1.272***
(β_4)	(0.0365)	(0.0366)	(0.0370)	(0.039)
Constant	14.57***	14.57***	14.89***	14.60***
	(0.0670)	(0.0673)	(0.116)	(0.242)
Observations	8458	8458	8458	8458
R^2	0.110	0.126	0.2514	0.3036
R^2_a	0.109	0.118	0.243	0.2066
rmse	2.606	2.593	2.402	2.46
ll	-20098	-20024	-19368	-19062

Robust standard errors, clustered at district level, in parentheses. * $p<0.05$; ** $p<0.01$; *** $p<0.001$***
DD-I estimates Equation 4.1. DD-II estimates Equation 4.2 without sector fixed effects. DD-III estimates Equation 4.2. DD-IV estimates Equation 4.3. There are unreported additional coefficients. DD-II has interaction dummies for each district (34) $\times$ each post-decentralization year (2). DD-III has, in addition to interaction dummies for each district and each post-dec year, sector dummies (17). DD-IV contains interaction dummies for each sector (17) $\times$ each district (34) $\times$ each post-dec year (2) - 1156 in total.

Table C.8: Treatment Heterogeneity across Sectors - I

VARIABLES	**Education**	**Health**	**Agriculture**	**Electrif.**	**Livestock**	**Revenue**
post1_treat1	**0.876**	**1.239**	**3.454***	**4.099*****	**4.438****	**0**
	(0.785)	(1.006)	(1.710)	(1.197)	(2.047)	(0)
post2_treat1	**3.461*****	**3.323*****	**7.631*****	**6.889*****	**10.75*****	**0**
	(0.827)	(0.916)	(1.388)	(1.097)	(1.231)	(0)
post1_treat2	**0.0414**	**0.131****	**-0.0600**	**0.224*****	**0.651*****	**0.268*****
	(0.0609)	(0.0524)	(0.111)	(0.0692)	(0.136)	(0.0659)
post2_treat2	**-0.145***	**-0.163****	**-0.0158**	**0.200**	**0.836*****	**-0.148**
	(0.0785)	(0.0669)	(0.126)	(0.146)	(0.138)	(0.192)
post1	0.0360	-0.0245	0.455***	-0.0126	-0.0667	-0.126***
	(0.0459)	(0.0328)	(0.165)	(0.0524)	(0.0421)	(0.0392)
post2	0.255***	0.111**	0.482**	0.174***	0.0728	-0.0420
	(0.0588)	(0.0533)	(0.183)	(0.0548)	(0.110)	(0.0652)
treat1	-2.428**	-1.971**	-6.148***	-4.183***	-12.54***	0
	(0.926)	(0.878)	(1.428)	(1.133)	(1.237)	(0)
treat2	-1.735***	-0.350***	-1.284***	-1.400***	-2.500***	-2.618***
	(0.0565)	(0.0518)	(0.126)	(0.117)	(0.0813)	(0.168)
Constant	15.89***	14.77***	14.47***	13.75***	16.01***	14.31***
	(0.0707)	(0.0696)	(0.174)	(0.0969)	(0.0781)	(0.0822)
Observations	1625	1579	575	696	234	397
R^2	0.112	0.035	0.282	0.216	0.747	0.315
rmse	2.702	2.147	2.364	2.583	1.696	1.919
ll	-3916	-3442	-1306	-1644	-451.1	-819.0

Robust clustered standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Each column represents estimation for DD-I specification in Table C.7 separately for each sector.

Some variables have been dropped due to collinearity. There are unreported additional coefficients.

Table C.9: Treatment Heterogeneity across Sectors II

VARIABLES	**Fisheries**	**Industry**	**Housing**	**Works**	**Adminis**	**Public health**
post1_treat1	**2.043**	**-0.412*** **	**8.899*** **	**1.513*** **	**4.600*** **	**0**
	(2.327)	(0.0393)	(3.145)	(0.542)	(1.631)	(0)
post2_treat1	**11.83*** **	**-1.395*** **	**8.322** **	**3.296*** **	**10.71*** **	**0**
	(0.339)	(0.0351)	(3.153)	(0.598)	(1.200)	(0)
post1_treat2	**1.424*** **	**0.341*** **	**-0.570** **	**0.344*** **	**-0.920*** **	**0.267** **
	(0.198)	(0.0607)	(0.224)	(0.0593)	(0.181)	(0.123)
post2_treat2	**1.273*** **	**0.0808**	**0.192**	**0.211**	**-0.296**	**-0.827**
	(0.197)	(0.132)	(0.315)	(0.134)	(0.201)	(0.553)
post1	-0.175***	-0.147***	-0.0357	-0.232***	1.787***	-0.0937
	(0.0626)	(0.0393)	(0.211)	(0.0570)	(0.172)	(0.0576)
post2	-0.115	-0.0244	-0.986***	-0.191***	1.713***	-0.0793
	(0.122)	(0.0351)	(0.0965)	(0.0494)	(0.179)	(0.0451)
treat1	-10.84***	4.215***	-5.374*	-0.593	-8.990***	0
	(0.0674)	(0.109)	(3.139)	(0.619)	(1.065)	(0)
treat2	-2.309***	-1.128***	-1.130***	-0.617***	0.179	-0.113
	(0.187)	(0.119)	(0.315)	(0.114)	(0.183)	(0.590)
Constant	13.02***	12.87***	14.66***	14.98***	13.53***	13.37***
	(0.0674)	(0.109)	(0.0776)	(0.0905)	(0.171)	(0.115)
Observations	220	205	77	1012	524	84
R^2	0.802	0.568	0.554	0.081	0.505	0.083
rmse	1.108	0.626	1.114	3.033	2.565	1.265
ll	-330.2	-190.4	-112.8	-2554	-1233	-135.9

Robust clustered standard errors in parentheses. *** $p<0.01$, ** $p<0.05$, * $p<0.1$

Each column represents estimation for DD-I specification in Table C.7 separately for each sector.

Some variables have been dropped due to collinearity. There are unreported additional coefficients.

Table C.10: Treatment Heterogeneity across Sectors (with Fixed Effects) I

VARIABLES	**Education**	**Health**	**Agriculture**	**Electrif.**	**Livestock**	**Revenue**
post1_treat1	**0.939**	**1.537**	**3.309***	**4.056*****	**4.288***	**0**
	(0.802)	(1.027)	(1.810)	(1.247)	(2.441)	(0)
post2_treat1	**3.404*****	**3.163*****	**7.537*****	**6.814*****	**10.41*****	**0**
	(0.850)	(0.935)	(1.477)	(1.179)	(1.479)	(0)
post1_treat2	**0.0451**	**0.129****	**-0.0600**	**0.209*****	**0.651*****	**0.268*****
	(0.0625)	(0.0539)	(0.118)	(0.0745)	(0.162)	(0.0723)
post2_treat2	**-0.151***	**-0.165****	**-0.00383**	**0.183**	**0.836*****	**-0.133**
	(0.0799)	(0.0688)	(0.132)	(0.154)	(0.164)	(0.213)
post1	-0.230	-5.569***	1.003***	-1.694***	0.392***	-0.353***
	(0.161)	(0.836)	(0.200)	(0.248)	(0.0871)	(0.127)
post2	0.528***	0.784***	0.339*	-0.390***	0.567***	0.659***
	(0.0703)	(0.0647)	(0.190)	(0.112)	(0.0932)	(0.0941)
treat1	-2.428**	-1.971**	-6.148***	-4.183***	-12.54***	0
	(0.945)	(0.898)	(1.519)	(1.192)	(1.471)	(0)
treat2	-1.735***	-0.350***	-1.284***	-1.400***	-2.500***	-2.618***
	(0.0577)	(0.0529)	(0.134)	(0.123)	(0.0967)	(0.184)
Constant	15.89***	14.77***	14.47***	13.75***	16.01***	14.31***
	(0.0722)	(0.0711)	(0.185)	(0.102)	(0.0929)	(0.0901)
Observations	1625	1579	575	696	234	397
R^2	0.128	0.075	0.340	0.280	0.817	0.367
rmse	2.733	2.147	2.410	2.603	1.717	2.022
ll	-3901	-3409	-1282	-1614	-413.3	-803.2

Robust clustered standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Each column represents estimation for DD-II specification in Table C.7 separately for each sector.

Some variables have been dropped due to collinearity. There are unreported additional coefficients.

Table C.11: Treatment Heterogeneity across Sectors (with Fixed Effects) II

VARIABLES	**Fisheries**	**Industry**	**Housing**	**Works**	**Adminis**	**Public health**
post1_treat1	**2.056**	**-1.169*** **	**8.664***	**1.437****	**4.719*****	**0**
	(2.749)	(0.114)	(4.314)	(0.570)	(1.724)	(0)
post2_treat1	**11.87*****	**-2.562*****	**7.935***	**3.295*****	**10.65*****	**0**
	(0.268)	(0.115)	(4.324)	(0.624)	(1.293)	(0)
post1_treat2	**1.424*****	**0.341*****	**-0.570***	**0.297*****	**-0.920*****	**0.267**
	(0.239)	(0.0742)	(0.308)	(0.0822)	(0.194)	(0.151)
post2_treat2	**1.273*****	**0.0808**	**0.192**	**0.204**	**-0.296**	**-0.827**
	(0.237)	(0.162)	(0.433)	(0.141)	(0.216)	(0.677)
post1	-0.122	0.288**	0.199***	0.433*	4.674***	-1.017**
	(0.141)	(0.114)	(0.0447)	(0.221)	(0.534)	(0.362)
post2	0.00608	-0.393***	-1.210***	-0.797***	2.999***	-0.261*
	(0.101)	(0.115)	(0.112)	(0.0879)	(0.191)	(0.133)
treat1	-10.84***	4.215***	-5.374	-0.593	-8.990***	0
	(0.0813)	(0.134)	(4.315)	(0.640)	(1.140)	(0)
treat2	-2.309***	-1.128***	-1.130**	-0.617***	0.179	-0.113
	(0.226)	(0.146)	(0.432)	(0.118)	(0.196)	(0.723)
Constant	13.02***	12.87***	14.66***	14.98***	13.53***	13.37***
	(0.0813)	(0.134)	(0.107)	(0.0937)	(0.183)	(0.141)
Observations	220	205	77	1012	524	84
R^2	0.876	0.782	0.581	0.135	0.570	0.325
rmse	1.058	0.545	1.483	3.044	2.560	1.330
ll	-278.6	-120.5	-110.3	-2523	-1196	-123.0

Robust clustered standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Each column represents estimation for DD-II specification in Table C.7 separately for each sector.

Some variables have been dropped due to collinearity. There are unreported additional coefficients.

Table C.12: D-D for Non-Decentralized Sectors (with Fixed Effects)

VARIABLES	**Health**	**Agriculture**	**Livestock**	**Police**
post1_treat1	**-0.195**	**2.352****	**0.167**	**6.206*****
	(0.226)	(1.044)	(0.164)	(1.665)
post2_treat1	**-0.403**	**-0.611**	**1.390**	**14.07*****
	(0.568)	(1.397)	(0.828)	(1.039)
post1_treat2	**0.549**	**0.0995**	**0.497****	**0.556*****
	(0.665)	(0.0709)	(0.213)	(0.137)
post2_treat2	**0.556**	**-0.163**	**0.528****	**0.516*****
	(0.703)	(0.353)	(0.207)	(0.117)
treat1	0.986	-9.282***	-9.562***	-16.17***
	(0.733)	(1.417)	(1.391)	(0.997)
treat2	1.401	-1.382***	-0.791***	-1.381***
	(0.876)	(0.131)	(0.253)	(0.115)
Constant	1.467	12.42***	10.47***	18.32***
	(0.952)	(1.038)	(1.383)	(0.163)
Observations	306	306	306	306
R^2	0.539	0.619	0.648	0.810
rmse	4.817	4.786	5.050	3.135
ll	-872.3	-870.3	-886.7	-740.9

Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Each column represents estimation for DD-II specification in Table C.7 for a specific sector.

There are unreported additional coefficients.

Table C.13: Difference-in-Difference-in-Difference (D-D-D) Results

VARIABLES	**DDD-I**	**DDD-II**	**DDD-III**	**DDD-IV**
post1_treat1_dev	**0.214**	**0.0859**	**0.0684**	**-0.990**
(β_{14})	(0.768)	(0.780)	(0.777)	(0.819)
post2_treat1_dev	**1.761*****	**1.639****	**1.644****	**1.416****
(β_{16})	(0.611)	(0.605)	(0.598)	(0.638)
post1_treat2_dev	**-0.262**	**-0.261**	**-0.262**	**-0.287**
(β_{15})	(0.174)	(0.175)	(0.175)	(0.186)
post2_treat2_dev	**-0.263**	**-0.268**	**-0.272**	**-0.273**
(β_{17})	(0.187)	(0.188)	(0.188)	(0.200)
post1_dev	0.923***	0.905***	0.907***	1.234***
(β_6)	(0.0442)	(0.0461)	(0.0460)	(0.166)
post2_dev	0.444***	0.435***	0.437***	1.549***
(β_7)	(0.0835)	(0.0853)	(0.0851)	(0.156)
treat1_dev	5.178***	5.178***	5.183***	5.224***
(β_{12})	(0.766)	(0.769)	(0.762)	(0.807)
treat2_dev	-0.705***	-0.705***	-0.703***	-0.705***
(β_{13})	(0.213)	(0.214)	(0.214)	(0.225)
post1_treat1	2.132***	2.132***	2.132***	2.132***
(β_8)	(0.391)	(0.392)	(0.392)	(0.413)
post2_treat1	3.611***	3.611***	3.611***	3.611***
(β_{10})	(0.406)	(0.407)	(0.407)	(0.429)
post1_treat2	0.425**	0.425**	0.425**	0.425**
(β_9)	(0.164)	(0.165)	(0.165)	(0.174)
post2_treat2	0.359*	0.359*	0.359*	0.359*
(β_{11})	(0.189)	(0.190)	(0.190)	(0.200)
dev	3.901***	3.901***	6.83***	3.901***
(β_5)	(0.546)	(0.548)	(0.678)	(0.577)
treat1	-8.507***	-8.507***	-8.507***	-8.507***
(β_3)	(0.603)	(0.605)	(0.605)	(0.640)
treat2	-0.538**	-0.538**	-0.538**	-0.538**
(β_4)	(0.214)	(0.215)	(0.215)	(0.227)
post1	-0.0762*	-1.920***	-1.828***	3.013***
(β_1)	(0.0388)	(0.156)	(0.162)	(0.322)
post2	0.0791	-0.671***	-0.690***	4.285***
(β_2)	(0.0844)	(0.154)	(0.157)	(0.327)
Observations	9682	9682	9667	9682
R^2	0.235	0.256	0.257	0.360

Robust s.e.'s in parentheses. *** p<0.01, ** p<0.05, * p<0.1.
All columns estimate Equation 4.4. DDD-II adds district × post_dec year dummies.
DDD-II adds sector dummies, DDD-IV adds sector × district × year dummies.
There are additional unreported coefficients.

Table C.14: Comparison of D-D for Decentralized and Non-Decentralized Sectors

	Agriculture		**Health**		**Livestock**	
VARIABLES	**District**	**Provincial**	**District**	**Provincial**	**District**	**Provincial**
post1_treat1	**3.313***	**2.352****	**1.539**	**-0.195**	**4.274***	**0.167**
	(1.831)	(1.044)	(1.040)	(0.226)	(2.483)	(0.164)
post2_treat1	**7.602*****	**-0.611**	**3.192*****	**-0.403**	**10.52*****	**1.390**
	(1.489)	(1.397)	(0.943)	(0.568)	(1.492)	(0.828)
post1_treat2	**-0.0600**	**0.0995**	**0.129****	**0.549**	**0.651*****	**0.497****
	(0.118)	(0.0709)	(0.0539)	(0.665)	(0.162)	(0.213)
post2_treat2	**-0.00382**	**-0.163**	**-0.165****	**0.556**	**0.836*****	**0.528****
	(0.132)	(0.353)	(0.0688)	(0.703)	(0.164)	(0.207)
treat1	-6.213***	-9.282***	-2.000**	0.986	-12.65***	-9.562***
	(1.532)	(1.417)	(0.905)	(0.733)	(1.485)	(1.391)
treat2	-1.284***	-1.382***	-0.350***	1.401	-2.500***	-0.791***
	(0.134)	(0.131)	(0.0529)	(0.876)	(0.0967)	(0.253)
Constant	12.41***	12.42***	12.71***	1.467	13.95***	10.47***
	(0.185)	(1.038)	(0.0711)	(0.952)	(0.0929)	(1.383)
Observations	575	306	1579	306	234	306
R^2	0.365	0.619	0.100	0.539	0.820	0.648
rmse	2.429	4.786	2.154	4.817	1.741	5.050
ll	-1286	-870.3	-3414	-872.3	-416.5	-886.7

Robust clustered standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Each column represents estimation for DD-II specification in Table C.7 separately for each sector.

There are unreported additional coefficients.

Table C.15: Probability of Reelection of District Nazim - Probit Estimation

VARIABLES	(1)	(2)	(3)	(4)
physical	**0.274****	**0.361*****	**0.321****	**0.330****
	(0.125)	(0.135)	(0.133)	(0.129)
social	**-0.375**	**-0.621***	**-0.588****	**-0.602****
	(0.234)	(0.322)	(0.281)	(0.276)
post1	0.0776	0.159	0.168	0.168
	(0.361)	(0.228)	(0.214)	(0.221)
density			8.84e-05	0.000264
			(0.000431)	(0.000735)
householdsize		-0.618	-0.651	-0.649
		(0.529)	(0.602)	(0.609)
pacca		1.294	0.711	0.794
		(1.160)	(1.663)	(1.696)
literacy			0.0151	0.0202
			(0.0438)	(0.0412)
diversity			0.658	0.848
			(1.587)	(1.674)
urban				-0.0111
				(0.0344)
Constant	-0.548**	2.962	2.683	2.547
	(0.270)	(3.795)	(4.650)	(4.618)
Observations	68	68	68	68
Pseudo R^2	0.093	0.174	0.181	0.184
ll	-40.06	-36.49	-36.16	-36.05

Standard errors in parentheses

*** $p<0.01$, ** $p<0.05$, * $p<0.1$