Identifying the Causal Effects of Social Capital in Labour Markets: Estimation Strategies and Empirical Evidence

YUNSONG CHEN

Nuffield College
Department of Sociology
University of Oxford

Trinity 2011



A Thesis Submitted in Partial Fulfilment
of the Requirements for the Degree of
Doctor of Philosophy in Sociology

Acknowledgments

Writing a thesis to get a doctorate takes years, but learning how to do better research can be a lifelong enterprise. The years in Oxford, gulfed by tears, toils, trials, frustration, as well as unfolding excitement and happiness, has not only opened to me a new door to the academia, but also has become the most precious experience in my life.

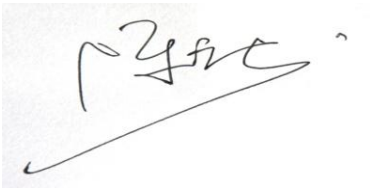
I would like to thank my academic supervisor, Professor Peter Hedström, for his inspiring, encouraging and very helpful supervision, invaluable gifts for which I feel so privileged. I believe that I will get more guidance from Peter in the future despite the vast distance between the Far East and Scandinavia. I am also indebted to Tom Snijders, Nan Dirk de Graaf, Jonathan Gershuny, Xiaogang Wu, Peilin Li, Steve Morgan, Colin Mills, Scott Moser, Tak Wing Chan, Anne Millard, Stephanie Wright, and Sandra González-Bailón. I feel fortunate for the opportunity that I had to learn so many things from them: they are the source of continual support as well as examples of intellectual strength.

Along the years, my fellows, Guiying (Laura) Wu, Zhiyun Li, Qianzi (Chloe) Zeng, Xiang Zhang, Yao Wang, Rui Wen, Jing (Becky) Yang, Suyu Liu, William Marks, and Aya Mastuo, have been a magic pool of camaraderie and friendship, and for that I am extremely grateful. My colleagues and friends in China also deserve my greatest thanks. They include Hao Ye, Zheng Song, Yu Bo, Jinsong Cao, Wei Wang, Xiaohong Zhou, and Xuewei Zhai. Especially, I owed a great personal debt to my mentor Hao Ye. Without his support I could not have started this journey, let alone have reached this point.

My family has been always an invaluable source of support and love. I am grateful to my wife Chun Hong for all the supports when I were off China to undertake this doctorate. I would like to give my thanks to my parents who taught me how to keep a humble heart armed with resilience and perseverance. To my daughter Jinya (Jenny), I often wonder how to compensate for your time without my presence, but you will know someday in the future that my heart has always been with you even though we had been seven seas apart.

I truly believe there is no culmination in such a journey, although it has been so far away from where I started. The very acknowledgment must have omitted many names who helped and are helping me to survive the rugged venture. However, the memory and gratefulness for all of them never fade away.

I declare that this thesis represents my own work, except where otherwise specified, and that it has not been previously included in a thesis, dissertation or report submitted to this University or to any institution for a degree, diploma or other qualifications.

Signature:  **Date:** 2011.9

Abstract

Identifying the Causal Effects of Social Capital in Labour Markets: Estimation Strategies and Empirical Evidence

Yunsong Chen
Nuffield College
Department of Sociology
University of Oxford

Trinity 2011

This thesis contributes to the sociological literature about how social capital influences individual labour market outcomes by using and combining a constellation of identification strategies to alleviate endogeneity bias. The principal proposition of social capital theory has been that social resources embedded in social networks have significant influences on individual labour market outcomes. Although a vast amount of empirical literature has accumulated, previous sociological studies of this line have not paid enough attention to the endogeneity problem rising from various sources such as omitted variables, self-selection, sample-selection, and simultaneity bias. Unless the endogeneity problem is explicitly pointed out and well-corrected for, one cannot arrive at any convincing causal conclusions about whether social capital matters or not.

Building on the existing literature, this thesis offers a systematic review in Chapter 2, paying close attention to the identification problem and empirical strategies. To identify causality under the counterfactual framework, four empirical chapters test for the effects of social capital using different identification strategies. Chapter 3 combines the instrumental variable method with the Heckman model to estimate network effects on labour market outcomes among rural-to-urban migrants in China. Chapter 4 presents evidence of the network effects on individuals' employment and income in China employing the instrumental variable method to control for the fact that friendships are not randomly formed. Using panel data and Heckman model from the former East Germany and contemporary China, Chapter 5 and Chapter 6 address the direct effects of using contacts to find jobs and the effects of contact's prestige in labour markets under state socialism.

Contents

Chapter 1: Introduction 8

Chapter 2: Empirics of Identification Challenges and Strategies: A Literature Review of the Role of Social Capital in the Labour Market 17

1. Identification Challenges and Strategies 18
2. Literature of the Role of Social Capital 47
3. Conclusion 74

Chapter 3: Network Effects among Rural-to-Urban Migrants in China 80

1. Introduction 80
2. Background 82
3. Models and Identification Issues 87
4. Data and Variables 96
5. Estimation Results 107
6. Interpretation of the IV Estimate 114
7. Conclusion 119

Appendix 1 121

Chapter 4: Do Friends' Diplomas Matter? Estimating Network Effects in the Presence of Selection 122

1. Introduction 122
2. Background 125
3. Models and Methods 131
4. Data and Variables 136
5. Results From Single-Equation Models 145
6. Results From the IV Models 148
7. Interpretation and Discussion 153
8. Conclusion 159

Appendix 1 161

Appendix 2 162

Chapter 5: The Effects of Using Contacts under State Socialism: Evidence from the First-Difference Analysis 163

1. Introduction 163
2. Literature Review 167
3. The First-Difference Model 171
4. Results from the GDR Data 173
5. Results from the Chinese Data 179
6. Further Discussion 188
7. Conclusion 195

Chapter 6: Do Contacts Matter under State Socialism? A Replication and Extension 197

1. Introduction 197
2. Background 200
3. Models 204
4. Evidence from the GDR 208
5. Evidence from China 215
6. Conclusion 224

Chapter 7: Conclusion 226

1. Empirical Findings 226
2. Limitations 231
3. Extension 232

References 235

Word count approximately 60,000

Chapter 1

Introduction

The social capital literature has long been criticized for definitional ambiguity (see e.g. Portes 1998, Durlauf 1999, Manski 2000). This thesis adopts the dentition of social capital by Lin (2001:29): “the resources embedded in social structure which are accessed and/or mobilized in purposive actions”. Since “embedded resources” is a general definition, it is compatible with Mouw’s (2006:80) proposition that social interactions such as peer effects or neighbourhood effects can be “subsumed under the definition of social capital”. In this regard, the characteristics, behaviours or outcomes of one’s reference group or a specific contact that influence the individual outcomes are all regarded as embedded resources[1]. Adopting such a broad concept of social capital effectively provides us with a richer literature. The empirics of the identification strategies for causal inference found in the wider literature can provide help on how to deal with the endogeneity problem that plagues the sociological analysis of social capital[2].

1 Similar definitions of this class of social capital are also provided by other sociologists and economists. For example, Durlauf (1999:2) defines social capital as "the influence which the characteristics and behaviors of one's reference groups has on one's assessments of alternative courses of behavior”. Portes (1998) offers a definition as “the ability of actors to secure benefits by virtue of their membership in social networks or other social structures.”

2 The causality this thesis refers to is defined in the counterfactual framework (see, for example, Holland 1986, Winship and Morgan 1999:702; Sobel 1996; Morgan and Winship 2007).

According to different operational definitions of social capital, the extensive literature can effectively be classified into three strands (Lin 1999, 2001; Mouw 2003). (1) "Using Contacts" as social capital. This line of research primarily addresses the treatment effects of using contacts or social networks on individual labour market outcomes. (2) "Used Social Capital". This points to the quantity or quality of social resources already mobilized or used from a certain helper(s). A typical research interest of this strand is to probe whether the contacts' resources (e.g. occupational prestige or income) matters for the respondents' relevant job outcomes. (3) "Accessed Social Capital" or "Network Social Capital". This strand of literature is mainly concerned with the quantity or quality of the social resources to which the respondent has access. Accessed social capital can be measured from two dimensions. One is the direct resource measure (e.g. the attributes of the job contacts or the mean group attributes of the networks) and the other is network structural measure (e.g. the size of network, the density or range of network or the structural hole).

The principal proposition of social capital theory, in the context of labour markets, has been that embedded resources have significant influences on individual labour market outcomes (e.g. labour force participation, labour migration behaviour, wage, occupational prestige, job satisfaction, working hours, promotion or performance at work, to name a few). Over the last few decades, a vast amount of empirical literature on social capital and labour market outcomes

has accumulated (see Granovetter 1995, Lin 1999, 2001, 2008 and Mouw 2003 for a review of sociological studies; see Angrist and Krueger 1999, Ioannides and Loury 2004 for a survey of economic literature)[3]. Although following different research traditions and methodologies, both sociologists and economists have long been concerned with the role of social capital in the labour market.

Since most of the empirical studies in sociology are based on non-experimental data, they naturally suffer from the endogeneity problem arising from a constellation of sources, such as omitted variable bias, self-selection bias, sample-selection bias and simultaneity bias[4]. However, in sociological studies on social capital in sociology, empirical studies paying close attention to the endogeneity problem are very rare. Mouw (2003) critically reassesses the previous sociological empirical findings on the causal effects of social capital in the job-matching process, arguing that self-selection in the forming of friendships (i.e. individuals tend to make friends with those similar to them) may at least account for part of the statistical association between social capital variables and individual job outcomes. Although Mouw (2003) is primarily concerned with the presence of self-selection as the major source of estimation bias, his work invites further attention on how to systematically deal with the endogeneity problem in estimating social capital's causal effects. Moreover,

3 There is also a rich literature on theoretical models of social capital (see e.g. Montgomery 1991, Topa 2001.Calvó-Armengol and Zenou 2005; see also Ioannides and Loury 2004 for a review).

4 Self-selection bias can also be seen as a special form of omitted variable problem. Meanwhile, both self-selection and sample-selection are called selection bias. In accessed social capital studies, self-selection is often termed an endogenous group membership problem (Mouw 2006; Moffit 2001).

Mouw (2006) offers a cogent review on how to deal with the endogeneity problem in estimating the effects of accessed social capital. However, estimation bias and identification strategies in empirical studies on the effects of the other two types of social capital (i.e., using contacts and used social capital) were not discussed in his work.

So, what are the substantive implications of the endogeneity problem for estimating the role of social capital in the labour market? Is it just a trivial issue or should it actually be a fundamental concern? Is it pervasive or confined to only a fraction of the existing studies? First, as Mouw (2003) argues, if social capital has little overall influence on labour market outcomes, then economic models of the labour market can safely ignore social ties among individuals. However, a fundamental problem here rises as, without explanatory power, the concept and perspective of social capital will degenerate into "only a lesson in the ambiguity of words" (Manski 2001:123). If the research interest lies in identifying a mechanism or causality, rather than merely describing the statistical association between the dependent and independent variables, these sources of endogeneity biases should actually be taken very seriously in the sociological analysis of social capital.

Second, the endogeneity problem is pervasive and can actually be found in all the three lines of social capital studies mentioned earlier. For example, numerous

pieces of research have shown that the factors determining people's job-seeking methods (e.g. job types, class, human capital, etc.) are closely correlated with job outcomes (Marsden 2001, Holzer 1987, Bentolia, Michelacci and Suarez 2010, Lin and Ao 2008). If these factors and the relevant unobservables are not controlled for, the estimated treatment effects of using contacts might be spurious (Mouw 2003). When it comes to identifying the effects of used social capital, aside from the conventional omitted variable problem, sample selection is an issue because we only look at those who used contacts to find work (Marsden and Hurlbert 1988, Lin 2003). As for the accessed social capital strand, due to selective friendship (for a review see McPherson, Smith-Lovin and Cook 2001), the "genuine" cause of the variation of the individual job outcomes might in fact be social homophily rather than network resources (Mouw 2003, 2006). Another problem is the simultaneity bias, as there is a mutual influence between the attributes of the reference group and the outcomes of the group members. In the existing sociological literature on the role of social capital in the labour market, however, limited attention has been paid to these problems. As a result, in spite of a large amount of empirical findings, so far the problem is still open: does social capital matter in labour markets?

Beyond labour market research, concerns about the endogeneity problem and the related correcting methods are not new. Although variously dubbed 'peer effects', 'neighbourhood effects', 'social-network effects', 'social contagion', 'social

diffusion' and 'learning' in different research contexts, the effects of social interaction in education, risk behaviour, migration and innovation adoptions have been widely explored. Many of the empirical studies have extensively exploited different statistical models, smart research designs and rich data sets to address the identification challenges. Examples of advanced models include instrument variables, the fixed-effects model, sibling effects, propensity score matching, experiments or quasi-experiments (see Angrist and Krueger 1991, Calvó-Armengol and Ioannides 2005, Mouw 2006, Soetevent 2006, Hartmann et al. 2008 for reviews).

Building on the existing literature, this thesis aims to review and apply various identification strategies to estimate the effects of social capital in labour markets under a variety of social contexts. It offers an overall review of the social capital literature, focussing on the pros and cons of the various identification strategies. It presents a set of empirical applications that exploit different identification strategies and rich datasets to help identify the causal effects of social capital.

Chapter 2 offers a systematic literature review of the broad concept of social capital, paying close attention to the identification problem and empirical strategies. Four major sources of endogeneity[5], namely, omitted variable bias, self-selection bias, sample-selection bias and simultaneity bias, are summarized

5 There are, of course, other problems hindering the identification of key model parameters (e.g. measurement errors and data collection strategies). These are beyond the scope of this thesis.

and the relevant identification strategies in previous studies are reviewed. This chapter revisits roughly 70 empirical studies.

Chapter 3 applies the instrumental variable (IV) approach in combination with the Heckman method to estimate network effects on labour market outcomes among rural-to-urban migrants in China. Rural-to-urban migrants in China return home frequently within a year. As a result, villages of origin can serve as an important intermediary where job leads can be transmitted among them. This pattern of return implies the pivotal role of the village-level outflow of migrants in determining their job outcomes at the destination. Using data from 22 provinces in China, this chapter assesses whether a migrant's wage is positively influenced by the number of fellow villagers who migrated to urban areas. The results confirm the significant effects of origin-based networks, and show that network effects are underestimated without taking the endogeneity into account.

Chapter 4 examines network effects on individuals' employment and income, controlling for the fact that friendships are not randomly formed. Data from the 2003 Chinese General Social Survey are analysed. To correct for the selection bias, the average network resource of a set of observationally identical individuals is used to instrument for one's network resource. The empirical results support the significant role of network resources, and strongly suggest the presence of social heterophily in friendship formation.

Chapter 5 addresses the direct effects of using contacts to find jobs on individual labour market outcomes as compared to other job-seeking methods. Data sets from the former East Germany (GDR) and from contemporary China are analysed. Since people may select job-search methods, the first-difference model is fitted to alleviate potential selection bias in estimation. Findings here show that there is little evidence supporting the direct effects of using contacts in socialist countries, consistent with previous findings in market economies. The results also reveal the different selection mechanisms through which people choose their job-seeking methods in China and the GDR.

Chapter 6 examines the effect that the mobilized social capital has on individual labour market outcomes under state socialism. Once again, data sets from the GDR and from contemporary China are analysed. To deal with the endogeneity problem, results from the standard ordinary least squares (OLS) model, the Heckman two-stage model and the first-difference model are compared. The GDR evidence demonstrates that contacts' status matters for the attained status of respondents, which verifies the finding in Völker and Flap (1999) using the same data. Results from the Chinese data indicate that contacts' membership of the Communist Party influences the wage of respondents.

In general, the thesis furthers the sociological analysis of the role of social capital on labour market outcomes in two respects. First, this is, to the knowledge of the author, the first attempt to give a systematic account of the progress made by social scientists in improving causal inferences on the role of social capital in labour markets. The full typology of social capital is examined and the four most salient sources of endogeneity that limit casual inference are all addressed. In previous empirical studies or reviews, only some have been examined. Second, drawing on updated literature and exploiting rich datasets, this thesis employs a constellation of identification strategies. In particular, the IV approach, Heckman two-stage method and the first-difference model are used and combined to alleviate the estimation bias. Many of these are here applied for the first time in the sociological analysis of social capital.

Chapter 2

Empirics of Identification Challenges and Strategies: A Literature Review of the Role of Social Capital in the Labour Market

This chapter aims to offer a systematic review of the progress made by social scientists in alleviating the endogeneity bias and improving causal inferences on the effects of social capital in labour markets. It is organized as follows: Section 1 depicts the four major identification challenges. Since they are not specific to labour markets, the basic models as well as correction methods employed in parallel research fields are described briefly. Section 2 proceeds to present an overall review of the existing literature regarding the role of social capital in the labour market. Three lines of studies are surveyed in light of distinct social capital variables: (1) the decisions or behaviours involved in using contacts (using contacts); (2) the quantity of social resources mobilized from a certain tied helper (used social capital); and (3) the quantity of social resources that can be accessed in a given social network (accessed/network social capital). Finally, Section 3 concludes.

1. IDENTIFICATION CHALLENGES AND STRATEGIES

When the goal of models is to consistently estimate the coefficients of interest for causal inference, a vital precondition is that the structures of alternative model specifications should be ruled out (Cameron and Trivedi 2005:20-30). Less formally, we can understand identifying a structure as eliminating the possibility of the existence of a competing explanation. Following Angrist and Krueger (1999:8), the identification strategy can be seen as a combination of "a clearly labeled source of identifying variation in a causal variable and the use of a particular econometric technique to exploit this information".

There are many potential ways in which an identifying model can be jeopardized (e.g. mis-specified functions, measurement errors and the endogeneity problem). Of interest to this chapter are the four major sources of endogeneity problem, because they are pervasive in identifying the causal effects of social capital in the labour market. These identification challenges are:

- Omitted variable bias;
- Self-selection bias;
- Sample-selection bias; and
- Simultaneity bias.

As a matter of fact, none of these four challenges is specific to research on the role of social capital in the labour market. Instead, they are all part of the canon of major identification problems in applied econometrics and empirical sociology[6]. What follows is a brief survey explaining why these identification problems matter and how they might be redressed using examples from labour economics and sociological literature.

1.1. Omitted Variable Bias

The omitted variable bias is dubbed the 'unobservable bias', 'hidden bias', 'confounding problem', or 'unobserved heterogeneity' in different disciplines. Consider a typical social capital model expressed as

$$y = \beta_0 + \beta_1 x_1 + \beta X + \varepsilon \qquad (1)$$

where y is the job outcome, x_1 is the social capital variable, the majuscule X denotes the other exogenous observed controls and ε denotes the error term. If ε is not correlated with x_1, we can get an unbiased estimate of β_1 from Eq. (1) with the OLS method. However, if there is an important omitted variable x_2 both varying with x_1 and y, the error term ε must contain x_2 and will be correlated

6 See Wooldridge (3rd ed. 2006, 2002) and Cameron and Trivedi (2005) for a fuller discussion and model derivation.

with x_1. As a result, the OLS estimate of β_1 will be biased and inconsistent because the Gauss-Markov Assumptions are violated.

How much can the magnitude of bias be? Suppose the 'true' model should be

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta X + \varepsilon \qquad (2)$$

and the omitted variable x_2 is correlated with social capital variable x_1 by

$$x_2 = \delta x_1 + v \qquad (3)$$

we use the symbol "~" rather than "^" to emphasize that $\tilde{\beta}_1$ is obtained from Eq.(1) (which is the underspecified model), and $\hat{\beta}_1$ and $\hat{\beta}_2$ come from Eq. (2). Thus, we get $E(\tilde{\beta}_1) = E(\hat{\beta}_1 + \hat{\beta}_2\tilde{\delta}) = \beta_1 + \beta_2\tilde{\delta}$. That is, the magnitude of the omitted variable bias in estimating Eq.(1) would be $\beta_2\tilde{\delta}$.

The omitted variable bias is of potential importance for all the three strands of social capital literature. Consider three typical wage models and unmeasured individual linguistic skill as a confounder. For the model estimating the treatment effect of using contacts, the estimated influence of using contacts on wage might be biased if the individual's linguistic skill is omitted from the specification. The reason is straightforward: linguistic skill may not only affect the opportunity to

use contacts, but also influence the kind of job and its wage because higher linguistic skill in itself represents higher ability. As for analysing the effects of the used social capital, one may have good reason to suspect that a job seeker with better linguistic skill actually has a higher probability of getting help from a more resourceful job helper. Thus, again, when the linguistic skill is omitted, the relationship between the job helper's occupational prestige and the respondent's wage can be spurious. When it comes to the role of accessed or network social capital on individual income, the estimated causal effects might well be challenged, since it is reasonable to expect that the linguistic skill can be positively related to the mean status or income of friends due to social homophily.

Having said that, in empirical studies there are four popular methods for tackling the omitted variable bias. First, one can exploit more and untraditional data to proxy and thus control for the potential omitted variables. For instance, in neighbourhood studies, Jencks and Mayer (1990) show that the magnitude of the estimated effects of growing up in a poor neighbourhood tends to decline as more controls for parental characteristics are included. Likewise, Duncan, Brooks-Gunn and Klebanov (1994) include family warmth and home learning environment as controls. In a school effects study, Clark and Lohéac (2005) control for a dummy reflecting whether parents selected neighbourhoods to proxy whether parents selected schools for children. However, it is never clear

whether all the unmeasured factors are controlled for by using a proxy variable. Instead of striving to build an exhaustive method (which is of course theoretically impossible), an alternative method is to use a lagged dependent variable as a proxy for the unobserved individual heterogeneity and historical factors. As Wooldrige (3rd ed. 2006:315) points out, this approach increases the data requirement for a cross-sectional study, yet it provides "a simple way to account for the historical factors that cause the current differences in the dependent variable that are difficult to account for in other ways".

The second approach relies on time variation permitting the use of a fixed-effects model to control for time-invariant unobservables. In essence, the comparison between the over time observations of the same peer group are used as the 'within variations' to help identify the model (see Plotnick and Hoffman 1995 for a survey on using the fixed-effect model to estimate neighbourhood effects; see also Halaby 2004 for a review on the panel analysis in sociology). In studies of the peer effect in schools, for example, Hanushek et al. (2003) use the panel Texas school data to control for fixed individual, school-by-grade and school effects, reporting that school-by-grade peer achievement has a positive impact on individual achievement growth. In fertility behaviour studies, Munshi and Myaux (2006) use panel data in Bangladesh and controlled for time-fixed individual heterogeneity. In social contagion studies, Ichino and Maggi (2000) account for individual and bank branch fixed effects to examine how a bank employee's

shirking behaviour is influenced by the mean group behaviour in bank branches.

In empirical studies, “genuine” panel data that have repeated observations for the same respondent are often difficult to collect. Nevertheless, under certain assumptions, the administrative records of organizations can offer opportunities for tracking down changes over time and thus allow for analysis of fixed effects. Exploiting the richness of the Texas Schools Microdata Sample, for instance, Hoxby (2000) uses the within-school year-to-year class variation (i.e. comparisons between classes in the same grade across adjacent cohorts and years) to estimate the role of classroom demographic composition on individual test scores. The estimates are argued to be unbiased under the assumption that the annual variation is random. In much the same way, Gould, Lavy and Paserman (2005) exploit the over time idiosyncratic variation in the number of immigrants within the same grade to identify the causal effect that the share of immigrant children in a grade has on students’ high school outcomes in Israel. Likewise, Lavy, Paserman and Schlosser (2008) use the idiosyncratic variation in the class proportion of repeaters-skippers-regulars to help identify the influence of class composition on regular students’ academic performance. Although strictly speaking these studies cannot be seen as “pure” panel analyses, they do exploit the “variation over time” of peer characteristics assumed to be random.

The third methodology to addresses the omitted variable bias, especially

group-level heterogeneity, is to exploit 'non-time within-group variation' in order to estimate fixed effects models. This is, in effect, a variation of the second method and has an identical identification strategy: to exploit more than one observation per group to conduct within-group analysis. Group-level heterogeneity is an issue in many social capital studies. For instance, a problem long plaguing the estimation of classroom effects is unobserved school characteristics, which arise because parents may select schools. However, if classes within the same school are assumed to be formed roughly randomly with respect to these unobservables, the variance between classrooms within the same school can be used to evaluate the classroom effects. Ammermueller and Pischke (2006) use within-school variation to identify the classroom peer effects on test scores in primary school in six European countries. Similar strategies can be found in McEwan (2003), estimating grade peer effects in secondary schools in Chile, and in Arcidiacono and Nicholson (2005) addressing whether the abilities and specialty preferences of a class affect a student's academic achievement and choice of speciality in a medical school. Aside from academic achievements, the with-school methodology has been used to probe the classroom effects on social ties behaviour (going out, expenditures, etc.) and risky behaviour by adolescents (see Soetevent and Kooreman 2005, Clark and Lohéac 2005).

This within-group strategy has been employed in neighbourhood effects studies

and social network effects studies to deal with the potential presence of the omitted household characteristic. The underlying logic of the strategy is that one can compare siblings who are raised in different neighbourhoods in order to evaluate the causal effects of their peers. Using the "family fixed effects" approach, Aaronson (1998) reports a significant and positive relationship between neighbourhood educational attainments and the individual propensity to complete high school. Based on the same data, Plotnick and Hoffman (1999) find no presence of neighbourhood effects on teen childbearing, high school graduation, and family income of females. Using the sibling data from the Mexican Migration Project, Palloni et al. (2001) confirm the effects of social capital on the decision to migrate. That said, as noted by Aaronson (1998), the fixed effects estimators suffer from large standard errors and do not control for time-varying unobserved factors.

The fourth approach addressing the omitted variable bias is the use of IVs – if one can find and convince readers that the IV candidates are not correlated with the unobserved factors (i.e. the error term) but strongly related to the explanatory variable. As noted before, family characteristics can always be an omitted variable in identifying neighbourhood, classroom or school peer effects. To get around this problem, the aggregate data at the state, county, or metropolitan area level are often exploited as IVs (see Card and Krueger 1996 for a review) to help identify the model. For instance, Evans, Oates and Schwab (1992) use the

metropolitan area unemployment rate, median family income and poverty rate to instrument the school effect variable: percentage of disadvantaged students. They report no significant school peer effect on teenage pregnancy or school dropout behaviour. Likewise, by instrumenting the segregation by local government fragmentation in each metropolitan area, as well as the number of rivers and streams flowing through each metropolitan area, Culter and Glaeser (1997) find that blacks in more segregated areas have significantly worse outcomes than blacks in less segregated areas. Similar applications of the IV approach in estimating peer effects can also be found in Duncan, Connell and Klebanov (1997) and Currie and Yelowitz (2000). However, as with fixed effect estimators, IV estimators tend to have large standard errors. More importantly, whether the IV is uncorrelated with the error term is untestable.

1.2 Self-Selection Bias

As a specific type of omitted variable bias, the self-selection problem refers to when the outcome/dependent variable of interest is "determined in part by individual choice of whether or not to participate in the activity of interest" (Cameron and Trivedi 2005:546). For instance, one may decide whether to attend a training course and the decision is made due to individual-specific factors. If gregarious people are more likely to attend, evaluating the causal effects of the training course on labour market outcomes may risk using a sample

in which gregarious people are over represented. In the context of labour markets, to understand the "selection" one could regard the process during which social capital exerts its effects as two distinct processes. One is sorting (e.g. selecting a certain job-seeking method, selecting friends, sorting to neighbourhoods, selecting schools) and the other is the process during which the social capital variable directly influences the job outcome. Compared to the conventional omitted variable bias, the self-selection problem occurs when the two processes are correlated due to one or more unobservables.

This raises the question, "do people have selection behaviours in labour markets?" In fact, numerous studies have confirmed the presence of a selection process in various contexts. For instance, individuals of higher status, more years of schooling and richer human capital might tend to use contacts less often (Marsden 2001; Bentolia, Michelacci and Suarez 2010; Lin and Ao 2008), or more often (Boxman, De Graaf and Flap 1991). In addition, according to social homophily theory, social ties are more likely to link similar individuals in terms of race, gender, social class, religion, behaviours and values (see McPherson, Smith-Lovin and Cook 2001 for a review). Given the presence of selection behaviour, Mouw (2006) argues that, despite the intuitive appeal of the effects of social capital being strong, scepticism based on a social homophily perspective may always exert its own appeal.

As noted before, the social capital process can in fact be viewed as two distinct processes. As a result, two different models can be written respectively. The illustrative model estimating the effects of social capital is given by

$$y = \beta_0 + \beta_1 x_1 + \beta X + \varepsilon \qquad (4)$$

which is exactly the same as Eq.(1). When it comes to the selection model, one could define the utility (U) to an individual from sorting (herself to a certain social capital treatment) as a function of social capital variable x_1, a vector of observed individual characteristics X' (X and X'can be identical), and ε' which is the individual characteristics known by the individual herself yet undocumented by sociologists. Thus, one arrives at:

$$U = F(x_1, X', \varepsilon') + \eta \qquad (5)$$

Suppose a factor x_2 influencing both of the two processes is unobserved by sociologists and thereby omitted from the equations. Put differently, suppose x_2 enters both ε and ε'. Given Eq. (5) stipulates the association between ε' and x_1, it is obvious that there must be a relationship between x_2 and x_1. As a result, there will be a relationship between x_1 and ε. Thus, the estimated β_1 from Eq.(4) will be biased.

In the context of social capital, one may be interested in the treatment effects of using contacts (see Winship and Mare 1992 for a full discussion on the treatment effects model as an extension of selection bias). Defining the decision of whether to use or not to use contacts (x_1 as a dummy) as a function of the utility perceived by the individuals, Eq. (5) therefore turns into

$$x_1 = G(U) = G\big(F(X', \varepsilon', \eta)\big) \qquad (6)$$

Another leading example of self-selection in social capital study is so-called "endogenous group formation" (see Moffit 1998, 2001). The problem arises because individuals with similar characteristics may tend to form social groups with each other, therefore, the subsequent correlation between group characteristics/outcomes and individual outcomes may merely reflect these common underlying factors, rather than a causal effect from group peers (i.e. social capital effects). Moffit (2001) gives a brief social interaction model to illustrate this. Specifically, the model is written as

$$y_{ig} = \beta_0 + \beta_1 \bar{y}_{-ig} + \beta_2 \bar{x}_{-ig} + \beta X_{ig} + \varepsilon_{ig} \qquad (7)$$

$$U_{ig} = F\left(\bar{y}_{-ig}, \bar{x}'_{-ig}, X'_{ig}, \varepsilon'_{-ig}, \varepsilon'_{ig}\right) + \eta_{ig} \qquad (8)$$

where the subscript *g* indexes the group and *i* indexes the individual. In addition, *y* is the individual outcome and x denotes the individual characteristics. The

mean value with subscript *–i* denotes social capital or the group mean value except for individual *i*. Despite more variables, one can easily see from the above that if there is an overlap between ε and ε' for individuals, neither β_1 nor β_2 in Eq. (7) can be consistently estimated.

In empirical studies, it is very likely that there is an unobserved factor both entering the error terms of the selection process model and the substantial process model. An example helps to clarify the models above. In relation to Eq.(4) and Eq.(6), suppose being shy (x_2) is an unobserved factor and affects both the decision to use contacts (x_1) and the wage (y). Because Eq.(6) entails that being shy is correlated with using contacts to get a job (x_1), the error term of Eq (4) must be associated with x_1 as well. As a result, Eq.(4) yields a biased estimate of β_1. Likewise, when it comes to Eq.(7) and Eq.(8), being shy can make the individual i select similarly shy people as her reference group and thus enters ε'_{ig} and ε'_{-ig}. As a result, there must be a relation between being shy and $\bar{y}_{-ig}$. At the same time, shy individuals are more likely to have lower wages so that being shy is absorbed by ε_{ig}. Consequently, ε_{ig} must be correlated with the mean group wage as social capital $(\bar{y}_{-ig})$, thereby leading to biased estimates in Eq.(7).

Turning now to the solutions for potential self-selection problems, the most intuitive method is to ensure that selection is not present. By nature, random assignment to treatment is the most robust approach dealing with self-selection.

This is because random assignment ensures that there is no presence of any correlation between group characteristics or treatment participation and unobserved factors. Experimental and natural experimental design can therefore be ideal approaches, often conducted with a random assignment to the treatment.

In treatment effect studies, Duflo and Saez (2003) use an experiment to probe the role of information diffusion in an individual's decision to enroll in a Tax Deferred Account retirement plan within a university. Jacob (2004) exploits the Chicago Housing Authority's decision to demolish certain public housing projects and compared the standardized test scores and school enrollment of relocated individuals with those who were permitted to stay. There are sets of empirical studies exploiting the US Federal Moving to Opportunity (MTO) programme as a natural experiment to estimate the relevant treatment effects. For example, Kling, Ludwig and Katz (2004) find that relocating to lower-poverty communities reduces arrests among female youth for violent and property crimes compared to the control group. Duncan, Clark-Kauffman and Snell (2004) report substantial programme-based improvements in adults' perceptions of neighbourhood safety and in adults' mental health. Sanbonmatsu et al. (2006) examine the treatment effects of the MTO on reading and mathematics test scores of children in families offered vouchers for moving to new communities, but found no significant effects. Ludwig, Duncan and Pinkston (2005) state that offering families unrestricted

housing vouchers without additional counselling appears to have little effect on economic outcomes. Aside from the field experiments, lab experiments focussing on social interactions can also be found in Falk and Fischbacher (2002) and Cason and Mui (1998).

Aside from the treatment effects studies, experiments or natural experiments are also used to address the more fundamental question of how group characteristics influence individual outcomes. For example, Sacerdote (2001) exploits variation in peer groups generated by the random assignment of freshman roommates at Dartmouth and finds evidence of peer effects on academic effort, GPA, and fraternity membership. Zimmerman (2003) exploits the fact that first-year roommates at Williams College are assigned randomly with respect to academic ability, and analyses the effects of roommate's prior STA score on an individual's grades.

Similarly, Hoxby and Weingarth (2006) exploit the reassigning policy in the Wake County school district. Using panel data, they conclude that Wake County's numerous reassignments would mainly have affected achievement through the redistribution of lower- and higher-achieving peers. In addition, in social interaction studies, the existence of an exogenous shock (e.g. a natural experiment, a policy intervention or naturally occurring events) can be exploited to construct an exclusion restriction (i.e. an IV) to help identify a model. For

instance, Liebman, Katz and Kling (2004) examine the effects of neighbourhood poverty from census tracts on individual outcomes (e.g. risky behaviour, mental health and physical health by gender). To deal with the selection problem arising from the sorting of individuals across residential neighbourhoods, they rely on variation induced by the MTO experiment and instrumented for neighbourhood poverty using interactions of site dummies and treatment group dummies.

Experiments and natural experiments have pros and cons. On one hand, they form the best available research design to rule out self-selection and many other classes of confounding factors in order to isolate causal effects. On the other, it is often not feasible due to being costly and time-consuming. Another issue worth mentioning is that it is very likely that social experiments such as MTO by themselves contain selection biases. For example, the MTO project only aims at those who are under a certain poverty level, so most of the participants were Blacks and Latinos (Sampson 2008, Clampet-Lundquist and Massey 2008).

The second approach to deal with the self-selection bias is matching. This method is widely used in identifying treatment effects and can only be implemented under the assumption that self-selection is totally caused by the *observed* variables. Rosenbaum and Rubin (1983) first proposed this approach and termed it as the Propensity Score Method (PSM)[7]. The intuitive idea is that

7 The propensity score is defined as the conditional probability of receiving a treatment given pre-treatment individual characteristics. The major steps of conducting PSM are: 1) using logit or

the bias can be reduced when the comparison of outcomes is performed using treated and control subjects as similar as possible. Morgan (2001) studies the effect of Catholic schools on learning and used PSM estimates to supplement the estimates obtained by regression models. To probe the neighbourhood effects on high school dropout rates and teenage pregnancy, Harding (2003) employs PSM methods to compare two groups of children who have identical observed factors and different neighbourhoods during adolescence. He found that those in high-poverty neighbourhoods are more likely to drop out of high school and have a teenage pregnancy than those in low-poverty neighbourhoods. However, PSM relies on large samples and substantial group overlap. More importantly, it is based on unverifiable assumption that the treatment assignment is determined by a constellation of observed covariates, unobserved confounders cannot be controlled for[8].

The third method is to explicitly model the process of selection. For instance, Angrist and Lavy (1999) use a regression discontinuity design to model the probability of a class being split in an experiment triggering a class size treatment. They then examine the plausible causal relationship between class size and student test scores. Many other studies have used discrete choice

probit equations to model participating the treatment and then calculate the probability for each respondent as a propensity score; 2) finding for each participant a non-participant who has the most similar propensity score; 3) comparing the outcome indicators of the matched pairs and the difference is the individual gains resting on the treatment; 4) Calculate the mean of these individual gains to obtain the average overall gain.

8 Harding (2003) resorted to sensitivity analysis to support the PSM and concluded that "unobserved factors would have to be unreasonably strong to account for the associations between neighborhood and the outcomes."

models and other econometric methods to analyse neighbourhood sorting, schooling consumption and social learning decisions (e.g. Bayer, McMillan and Rueben 2002; Conley and Udry 2010; Van Der Klaauw and van Ours 2003). Finally, the Heckman selection estimator can also be used to address the self-selection bias by modelling the participation decision process. This method will be addressed below when this chapter deals with the sample-selection bias. However, it should be stated here that the limited application of Heckman's method in empirical studies may be due to the availability of a number of other approaches such as experiments, IV estimation and panel models (Bushway, Johnson and Slocum 2007).

Aside from the above three methods, one should note that, since self-selection can be seen as a specific type of omitted variable bias, those identification strategies for omitted variable bias can all be used in dealing with it. For example, including more controls or using proxies are often used as ways to deal with self-selection. For example, Hartmann and Yildiz (2006) deal with the endogenous group formation by including a rich specification for heterogeneity. Krauth (2004) looks at the peer effects on smoking among Canadian youth by using the degree of self-selection on observables as a proxy for the degree of selection on unobservables.

1.3 Sample-Selection Bias

Sample-selection bias arises when the dependent variable is observed only for a restricted, nonrandom sample. This nonrandom exclusion of certain observations can not only be caused by data collection procedures but also by processes inherent in the phenomena under study. An example is estimating the effects of used social capital. To probe the role of used social capital, an intuitive method is to fit a linear model onto those who reported a job helper and see whether the volume of the mobilized resources (e.g. the occupational prestige of the job helper) significantly affects the job outcomes. In a sample-selection model like this, however, common unobservables may affect both the individual job outcome and the probability of using contacts to find jobs in unknown ways. So, if we fit a simple OLS model onto the subsample of those who reported using contacts, this bias will occur.

Heckman (1976, 1979) formalizes the selection bias problem resulting from using nonrandomly selected samples in estimating behavioural relationships. He in fact regarded it as an ordinary specification error or omitted variable bias. Berk (1983) was among the first to introduce the selection bias in sociological literature. Here, a simple model is presented to clarify the basic issues. Suppose one is interested in the statistical properties of the OLS estimation of the population model

$$y = \beta_0 + \beta_1 x_1 + \beta X + \varepsilon \qquad (9)$$

Define a selection indicator *s* by *s*=1 if y are observed and s=0 otherwise. Rather than estimating Eq. (9), one is in fact only estimating

$$Sy = S\beta_0 + \beta_1 Sx_1 + \beta SX + S\varepsilon \quad (10)$$

The condition upon which Eq.(10) yields unbiased estimates regarding $S\beta_1$ is that $s\varepsilon$ should not be correlated with sx_1. Since *s* is a binary variable, the condition is in fact that ε should be uncorrelated with sx_1.

Now define *s* by

$$S = F(X', \varepsilon') + \eta \quad (11)$$

where X' is always observed and ε' is unobserved. F is a function that determines whether the selection indicator s is one or zero. It is natural to see that X' and X may overlap and so do ε' and ε. Importantly, if there is some unobserved or omitted factor, x_2 influences not only *y* in Eq.(10) but also *s* in Eq.(11), meaning the error term ε must be correlated with sx_1because it in itself contains x_2 and x_2 predicts *s*. That is to say, the estimates from Eq.(9) must be

biased if there is a certain unobserved factor affecting both the sample selection and the outcome. Note that Eq.(11) is similar to Eq.(5). This is why the sample-selection bias and self-selection bias are often both known as selection bias.

The most widely used approach to get around the sample-selection bias is the two-stage solution developed by Heckman (1976). Heckman uses a function to denote Eq. (11) by

$$s = 1 \text{ if } \gamma_0 + \gamma_1 x'_1 + \gamma_2 x'_2 + \cdots \gamma_m x'_m + \varepsilon' > 0 \qquad (12)$$

In the first step, using all observations, a probit model of s on X' is estimated and then the estimates $(\hat{\gamma}_0, \hat{\gamma}_1, \hat{\gamma}_2, \cdots \hat{\gamma}_m)$ are obtained. Following this, we compute the inverse Mills ratio (or hazard ratio) $\hat{\lambda}_i$ for each observation i. As Berk (1983) points out, the hazard ratio from a linear probability model is equal to the predicted probability of nonresponse minus 1.0, while that from a logit model is simply the predicted probability of nonresponse, namely s=0. The second step is then to use the selected sample (s=1) and fit the model where the independent variables include $(x_1, x_2, \cdots x_k, \hat{\lambda})$. By including the hazard ratio as a certain omitted variable in Eq.(9), one can consistently estimate β_1.[9]

9 As Wooldridge (3rd ed. 2006:620) mentions, to practically distinguish the sample selection from a mis-specified model, x should be a strict subset of X'. That is, any independent variables appearing in the substantial model, Eq.(9), should be included into the selection model, Eq.(11). Moreover, we should have at least one element of X' that does not have any partial effect on y.

Heckman's two-stage method is widely implemented in the social sciences due to its simplicity and the prevalence of OLS regression techniques in research (Hughes 1997). Still, it is far from watertight. In fact, critiques of this method are not new. Stolzenberg and Relles (1997) propose a method to judge whether the two-step estimator of Heckman is likely to increase or decrease the accuracy of regression coefficient estimates. Bushway, Johnson and Slocum (2007) highlight four common problems with its application, including the problematic use of dichotomous dependent variables[10], difficulties with calculating the inverse Mills ratio, misestimated standard errors, and co-linearity between the correction term and other regressors in the substantive model of interest.

Since sample selection can also be seen as a specific type of omitted variable problem, the IV method can actually be used to deal with it. For example, Kling et al. (2004), Kling, Liebman and Katz (2007), and Kling, Ludwig and Katz (2004) examine the treatment effect of the MTO project on those who did move using an MTO voucher, known as the effect of the treatment-on-treated (TOT) on various outcomes (crime rate, adult economic self-sufficiency and physical health). The sample-selection problem is an issue because some of those who received the voucher did not move, although the offering of the voucher to move is random.

10 For example, if we are interested in examining whether using contacts leads to job satisfaction or probability of employment which are generally operationalized as a dichotomous choice or other forms of discrete variables, it is problematic to resort to Heckman's two-stage method to alleviate the selection bias.

Thus, the survey sample must be a subpopulation of those who are eligible to move. If the decision to move is endogenous, estimating the TOT must be plagued by sample-selection bias. They thus use the indicator of assignment to MTO project or not (voucher offered or not) to serve as an IV for the treatment, i.e. using the voucher or not. The IV is valid because the assignment to the treatment group to get the voucher is random, so that it will not be related to unmeasured individual factors influencing individual outcomes, and it obviously is related to the decision to use the voucher or not.

1.4 Simultaneity Bias

An important form of endogeneity of explanatory variable is simultaneity. The simultaneity problem arises when the explanatory variable is jointly determined with the dependent variable. In social capital studies, it often occurs when there are mutual causal effects between the individual outcome variable and the social capital variable. For example, in accessed social capital studies, we may regress individual current occupational prestige score on the mean occupational prestige score of the respondents' social networks. However, due to the mean value obviously being affected by the individual outcome simultaneously, fitting a standard linear-in-means model yields biased estimates. This chapter draws on Manski (1993) and Moffit (2001) to give a brief illustration.

Without losing generality, a linear-in-means model estimating the effects of accessed social capital can be written as

$$y_{ig} = \beta_0 + \beta_1 \bar{y}_{-ig} + \beta_2 \bar{X}_{-ig} + \beta_3 X_{ig} + \varepsilon_{ig} \qquad (13)$$

where *i* indexes the individual, *g* indexes the group and x denotes the individual characteristics (e.g. gender, educational attainment). Note that $\bar{y}$ is the mean value of group outcome and the subscript *–i* shows that the mean value is calculated excluding the individual herself and $\bar{x}$ is the mean group characteristic. As a matter of fact, both β_1 and β_2 can be seen as the effects of social capital (Durlauf 2002).

To illustrate the problem, suppose there are only two individuals (i=1,2) per group, which means Eq. (13) can be written as:

$$y_{1g} = \beta_0 + \beta_1 y_{2g} + \beta_2 X_{2g} + \beta_3 X_{1g} + \varepsilon_{1g} \qquad (14)$$

$$y_{2g} = \beta_0 + \beta_1 y_{1g} + \beta_2 X_{1g} + \beta_3 X_{2g} + \varepsilon_{2g} \qquad (15)$$

If we plug the right-hand side of Eq.(14) in for y_{1g} in Eq.(15), we get the reduced form of Eq. (13) expressed by

$$y_{2g} = \beta_0 + \beta_1(\beta_0 + \beta_1 y_{2g} + \beta_2 X_{2g} + \beta_3 X_{1g} + \varepsilon_{1g}) + \beta_2 X_{1g} + \beta_3 X_{2g} + \varepsilon_{2g}$$

or

$$y_{2g} = \frac{\beta_0}{1-\beta_1} + \left(\frac{\beta_1\beta_2+\beta_3}{1-\beta_1^2}\right) X_{2g} + \left(\frac{\beta_1\beta_3+\beta_2}{1-\beta_1^2}\right) X_{1g} + \left(\frac{\beta_1\varepsilon_{1g}+\varepsilon_{2g}}{1-\beta_1^2}\right) \quad (16)$$

This thus leads to two problems: first, y_{2g} must be correlated with $\left(\frac{\beta_1\varepsilon_{1g}+\varepsilon_{2g}}{1-\beta_1^2}\right)$ because y_{2g} is correlated with y_{1g} and therefore ε_{1g}. As a result, estimates from Eq.(16) will be biased. This is called the simultaneity bias (Wooldridge 3rd ed. 2006). Second, only $\frac{\beta_0}{1-\beta_1}$, $\left(\frac{\beta_1\beta_2+\beta_3}{1-\beta_1^2}\right)$ and $\left(\frac{\beta_1\beta_3+\beta_2}{1-\beta_1^2}\right)$ can be identified, and one cannot separate β_1 from β_2. Manski (1993) terms this the "reflection problem". We first look at the approaches to deal with the reflection problem, and then turn to solutions for the simultaneity problem.

Manski (1993) is among the first to formalize the reflection problem in social interaction studies. According to him, individual outcomes are influenced by group characetistics through two different types of social effects: 1) endogenous effects denoting the influence of a certain group outcome on the same individual outcome ($\overline{y}$); and 2) exogenous or contextual effects pointing to the group characteristics' (rather than the outcomes') influence on individual outcomes ($\overline{x}$). In the labour market context, an example helps to clarify the difference. Consider a job seeker tied to a number of contacts in a given social network; all else being equal, there is an endogenous effect when her wage is predicted by the mean wage of her friends and there is exogenous effect when

her wage is influenced by the mean education level of her friends. Due to the reflection problem, it is often difficult to distinguish the exogenous and endogenous effects.

The reflection problem can be avoided by two simple methods.[11] First, avoiding using the linear-in-means model can be a solution. Manksi (2000) argues one can use the median or the subjective data of the group attributes rather than the mean value to serve as the major explanatory variable of interest. For example, using the unique data from the National Longitudinal Survey Youth (NLSY) 97, Kawaguchi (2004) uses teenagers' subjective perceptions of peer behaviours to identify peer effects in American youth substance use. Second, introducing non-linearity into a linear-in-means model would help identification (Manski 2000). Furthermore, Brock and Durlauf (2001) proved that the reflection problem can be avoided in the non-linearity model. This opens up possibilities, especially for research on the social capital effects on discrete dependent variables such as labour market participation (e.g. job satisfaction). Recently, Bramoullé, Djebbari and Fortin (2007) consider an extended version of the linear-in-means model and demonstrate that endogenous and exogenous effects can be generally identified if the size of group differs (i.e. respondents have

11 One situation is worth noting in which the reflection problem is likely to be minor in relation to the other sources of endogeneity. Sometimes, for policy purpose, it will not matter whether endogenous or exogenous effects give rise to social capital effects so that one of them may not be considered in the model which is still of substantive interest. For example, Ammermueller and Pischke (2006) do not distinguish the two effects and just report composite estimates of class effects on students.

different numbers of group/network members). Although this relaxes some conditions entailed in the model proposed by Manski (1993), the relaxed conditions are in effect closer to the real situation, such as individual-specific networks.

Before turning to the identification strategy for dealing with the simultaneity bias, we first examine some specific situations where one could plausibly assume that the simultaneity led by the mutual causality can be broken. For instance, Nair, Manchanda and Bhatia (2006) assume that the physician is influenced by the actions of his self-reported opinion leader (a high-prestige physician), but the opinion leader will not be influenced by the physician's actions. Likewise, Sorensen (2006) studies the social interactions inherent in the health-plan choices of university employees. He assumes that new hires are potentially influenced by the observed choices of existing employees, but not vice versa. In Kat, Kling and Liebman (2001), they assume that the incoming family does not influence the group variable of interest in the neighbourhood they move to. The reason for this is that, provided there is a sufficiently high number of existing families in the neighbourhood, the causal effect of social interaction is unidirectional.

In addition, one can impose a temporal ordering to break the simultaneity between the dependent and independent variables. Manski (1993) proposes that

this is in order "to exploit the recursive structure of equation". That is, one can assume that the focal individual's behaviour in time (t+1) is in fact affected by the group behaviour up to time (t). In empirical studies based on observational data, however, things are complicated. As Hartman et al. (2008) and Soetevent (2006) caution, a caveat to this identification strategy is that unobservables should not be correlated over time, and agents must be assumed not to be forward looking. For example, Hanushek et al. (2003) use the lagged peer achievements as a proxy for the current value, which is the explanatory variable. As they concede, this strategy will in general lead to an estimate that is downwardly biased if the error terms are serially correlated (e.g. because of unobserved teacher quality). Another problem with this method is that, according to Manski (1993), one must maintain the hypothesis that the transmission of social effects really follows the assumed temporal pattern. However, there is very little empirical research that provides evidence for any particular timing.[12]

In general, adding exclusion restrictions is the most accessible identification approach to solving the simultaneity problem. That is, one could find at least one exogenous variable which does not appear in Eq.(13) but appears in the model representing reverse causality. In the language of econometrics, only by doing so can we satisfy the rank and order condition for the identification of the social capital model (Durlauf 2003). In empirical studies, including an exclusion

12 That is, we do not have prior knowledge regarding how long the transmission of social capital takes.

restriction is equivalent to finding IVs for the relevant social capital variable. Recent studies often exploit the exogenous policy interventions or a certain random shock on the subpopulation of the peer group to serve as the source of the IV. This method, in essence, is identical to the "partial-population experiments" of Soetevent (2006) or the "preference changing intervention" coined by Moffit (2001).

For example, Boozer and Cacciola (2001) use the proportion of the class previously exposed to the "small class" treatment as an IV for current peer group achievements to identify the class effects on individual academic performance. This IV is ideal because the exposure is from a random design experiment, meaning it will not be related to individual or family heterogeneity. Meanwhile, the proportion of exposure to the treatment will obviously influence the current mean class academic achievements. In much the same way, Bobonis and Finan (2005) use the proportion of the village peers' participation in the Progresa programme in Mexico (involving payments to poor mothers)as an IV to examine how the school-enrollment decisions of village peers influence individual school-participation decisions. Some researchers use natural disasters such as earthquakes or environmental pollution as a source of random shock. Cipollone and Rosolia (2007) use data on earthquakes and the subsequent relief from military training of males to instrument class composition and identified peer effects on the school completion rate of Italian females. Asadullah and Chaudhury

(2008) identify endogenous social effects in mathematics test performance for eighth graders in rural Bangladesh using the proportion of peers exposed to arsenic contamination of water wells at home as an instrument.

2. LITERATURE OF THE ROLE OF SOCIAL CAPITAL

As noted earlier, there are three lines of research on social capital's labour market effects. Each line reflects the different social mechanisms through which the related social capital variable exerts its role. Focussing on estimation strategies to deal with the endogeneity problem, the following section presents an empirical treatment of the social capital studies in the context of labour markets addressing the three strands of the literature in turn.

2.1 Estimating the Treatment Effects of Using Contacts

Although there is no previous sociological research exclusively focussing on the direct impact of using contacts on labour market outcomes, there has been a research tradition probing the effects of the used social capital (see for a review Granovetter 1995, Lin 1999, Marsden and Gorman 2001, Mouw 2003). A summary of these empirical studies appears in Table 1. As we can see, although some of the literature (Campbell and Rosenfeld 1985, Simon and Warner 1992,

Peterson, Saporta and Seidel. 2000, Margolis and Simonnet 2003) supports a positive and significant association between using contacts and job outcomes, more studies report no significant effect or even a negative effect on job outcomes. Importantly, the findings espousing a positive role for using contacts are either based on small sample sizes or on a single firm, or point to specific contacts and show mitigated effects (see also Mouw 2003 for a review).

In the recent literature, several studies exploit the fixed effects model to rule out time-invariant unobservables to deal with the endogeneity problem. They all report that using contacts does not pay off. Mouw (2002) find conflicting evidence from cross-sectional and longitudinal data. He reports that although the OLS estimates indicate a negative relationship between using contacts and wages, the estimates from the fixed effects model show insignificant effects. In further research, Mouw (2003:880-81 Table 4) also fits a fixed effects model on panel data (NLSY 1994-1998) and finds that there is no significant association between using contacts and the log wages in US in the 1990s. Based on NLSY data from another time period (1981-1982) and using first-difference analysis, Kugler (2003) provides evidence that the positive and significant OLS and fixed-effects estimates of the effects of using employee referrals become insignificant once the employment sector is controlled for.

Table 1: Treatment Effects of Using Contacts in Previous Research (Dummy of using contacts as independent)

Research	Job outcome	Data and model	Findings
Lin, Vaugh & Ensel 1981 (USA)	Status first job	Two cross-section data sets OLS	None
Campbell & Rosenfeld 1985 (USA)	Wages	Cross-section OLS	Positive
Bridges & Villemez 1986 (USA)	Wages	Cross-section OLS	None
Wielgosz & Carpenter 1987 (USA)	Job satisfaction	Cross-section OLS	None
De Graaf & Flap 1988 (FRG, The Netherlands, USA)	Wages	Cross-section data	None
Marsden & Hurlbert 1988 (USA)	Wages	Cross-section OLS	None
Simon & Warner 1992 (USA)	Starting wage	Cross-section OLS	Positive
Vecchio 1995 (USA)	Job dissatisfaction	Cross-section Hierarchical regression	None
Korenman & Turner 1996 (USA)	Wages	Cross-section OLS	Positive (W) None (B)
Bian & Ang 1997:1002 (China & Singapore)	Status	Cross-section OLS	Positive (Cn) None(Sg)
Völker & Flap 1999: Table 5 (East Germany)	Wages/status of 1989 Status of first job	Panel OLS on each time period	None
Elliot 1999 (USA)	Wages	Cross-section OLS	None
Smith 2000 (USA)	Wages	Cross-section OLS	Negative
Peterson, Saporta & Seidel 2000 (USA)	First and final wages	Cross-section OLS	Positive
Falcon & Melendez 2001 (USA)	Wages Status	Cross-section OLS	None (B) Negative (B)
Mouw 2002 (USA)	Wages of black Americans	Panel OLS/fixed effects	None
Mouw 2003 (USA)	Wages/job satisfaction/ duration of employment	Panel OLS/fixed effects	None
Margolis & Simonnet 2003 (France)	Wages	Cross-section OLS	Positive
Kugler 2003 (USA)	Wages	Panel OLS/first difference	None

Loury 2006 (USA)	Job tenure Job satisfaction Wages and Wage growth	Cross-section OLS/probit	Mixed
Delattre & Sabatie 2007 (France)	Wages	Panel OLS/IV/switching regression	Negative
Mosca & Pastore 2008 (Italy)	Wages in various sectors	Cross-section OLS	Mixed
Ponzo & Scoppa 2009 (Italy)	Wages in private sector	Cross-section OLS	Negative
Bentolila, Michelacci & Suarez 2010 (USA, Europe)	Wages	Cross-section IV method	Negative

Notes:

Positive= Positive effects and significant at 0.05 level

Negative= Negative effects and significant at 0.05 level

None= Insignificant effects Mixed= Mixed effects depending on different dependants

W=White people B=Black people Cn=China Sg=Singapore

The IV method is also applied to estimate the treatment effects of using contacts to find work. Bentolia, Michelacci and Suarez (2010) use two IVs to deal with the endogeneity problem. One is the number of older siblings and the employment rate in the US is instrumented using contacts. The justification is that, the more older siblings one has, the higher the stock of network contacts one has and the higher probability of using contacts. At the same time, the sibling number is argued to be exogenous and has no direct role on worker productivity. The other IV is the employment rate in the US state that the worker lived in at 16 years old. This is because the most valuable job contacts are those who are employed. Thus, the "useful" contacts stock should be positively related to the local employment rate and therefore the probability of using contacts should be higher. Again, the state employment rate does not directly affect individual job outcomes (when the

relevant attributes of the state are controlled for). With IV estimates they contend that, due to mismatch between a worker's occupational comparative advantage and their actual productivity, using contacts in fact leads to lower wages. However, as noted before, the exogeneity of the IV cannot be directly tested, so that the finding is robust only when the readers believe that the IV is sufficiently valid.

To deal with possible selection bias, Delattre and Sabatie (2007:223) compare the results of three alternative methods (OLS, Heckman two-step method, and endogenous switching regression) to probe the role of mobilizing social networks on wages. The endogenous switching model is preferred since it allows for not only testing the network's choice endogeneity, but also whether wage explanatory factors are different according to the choice to use a network or not. For example, they find that, for those who found a job using network contacts, only the firm and job characteristics (e.g. contract duration, job skill level and firm size, each of them having a positive effect on wages) explain hourly wages, but individual attributes (e.g. age, race) have no impact except in the case of sex. According to their findings after correcting the selection bias, the effect on wages of mobilizing one's social network is negative.

2.2 Estimating the Effects of the Used Social Capital

Central to this line of research enterprises is Lin's (1999) social resource theory. That is, the more mobilized resources are, the better job outcomes will be. Leading empirical studies include Lin, Vaugh and Ensel (1981), Lin, Ensel and Vaugh (1981), Marsden and Hurlbert (1988), De Graaf and Flap (1988), Requena (1991); Wegener (1991), Bian and Ang (1997), Bian (1997), Völker and Flap (1999), Smith (2000), Mouw (2003), and Lin (2003). Table 2 summarizes these studies and shows that the most widely used measure of the used social capital is the occupational prestige of the job contacts. As for the job outcomes, status and wages are typically used.

As can be seen from Table 2, the majority of the existing studies demonstrated positive effects from used social capital, even though some of them are not statistically significant. This looks consistent with Lin's (2001) review stating that a consensus has been reached that high-status job helpers are beneficial to job outcomes. However, as Table 2 shows, all the studies are based on cross-sectional analysis and the OLS model. This means that the omitted variable and selection problems are still present.[13] In particular, the only effort to

13 Since most of the related survey questionnaires explicitly asked the respondents to provide information on a specific job helper (such as "what's the occupation of the job helper when he helped you to get current job?"), the time order is clear so that there is no mutual causality problem. That is, the simultaneity problem will not rise.

alleviate any potential endogeneity problem is merely including more controls (such as prior job outcomes) to control for unmeasured individual factors[14].

Table 2: Effects of the Used Social Capital in Previous Research (Contacts' prestige as independent)

Researches	Dependent	Data and Model	Finding
Lin, Vaugh & Ensel 1981 (USA)	Status first/current job	Two Cross-section data OLS	Positive
Lin, Ensel& Vaugh 1981 (USA)	Status first/current job Current income	Two Cross-section data OLS	Positive
Marsden & Hurlbert 1988 (USA)	Current Status Current Wages	Cross-section OLS (Heckman correction)	Positive None
De Graaf and Flap 1988 (FRG, Netherlands, USA)	Status first/current job Wages first/current job	Two Cross-section data OLS	Positive None
Requena 1991 (Spain)	Current income/status	Cross-section OLS	None
Wegener 1991 (FRG)	Current Status	Cross-section OLS (Heckman correction)	Positive
Bian & Ang 1997 (China and Singapore)	Current Status	Cross-section OLS	Positive
Bian 1997 (China)	Current Status	Cross-section OLS	Positive
Völker & Flap 1999:42 (East Germany)	Status of first/1989 job Wage of 1989 job	Panel OLS on each time period	Positive None
Smith 2000 (USA)	Wages	Cross-section OLS	Positive
Mouw 2003 (USA)	Current Status	Cross-section OLS	None
Lin 2003 (China)	Current Status (rank, prestige, wage)	Cross-section OLS (Heckman correction)	Positive

Notes:

Positive= Positive effects and significant at 0.05 level

Negative= Negative effects and significant at 0.05 level

None= Insignificant effects

14 As discussed before, this method is reasonable but far from perfect. This is because to proxy unobserved individual factor using job outcomes from a previous period (immediate prior job or the first job) introduces a number of additional unobservable factors into the model.

It should be noted that early studies such as Lin, Ensel and Vaugh (1981) considered the possibility that a person with a higher social position selectively connects to many other positions on the same level. However, they report no evidence that the higher a person's status, the higher the status of his or her contacts. Mouw (2003) was among the first to notice that social homophily might be a competing theory to the social resource perspective, i.e. that the social homophily factor may be an omitted explanatory variable. He thus replicates the study by Marsden and Hurlbert (1988) and extends it by including a dummy denoting the similarity of the occupation between the respondents and the job helpers. When the similarity of occupation is controlled for, the significant effects of used social capital disappear. He then concludes that 'the evidence in favor of the social resources perspective is largely an artifact of the incidence of same-occupation information flows between contacts and jobseekers' (Mouw 2003:883). Although Mouw (2003) operationalizes the social homophily factor (i.e. the similarity of occupation) in a problematic way, his argument that the used social capital might be endogenous is insightful. However, few sociological studies have furthered this issue to systematically deal with the potential endogeneity problem.

Aside from the omitted variable problem, sample selection is the other leading source of endogeneity affecting casual inference for the used social capital. For

example, individuals of higher social status tend to less use contacts (Marsden 2001). As a result, when one regresses the job status on the used social capital, using a "truncated sample" leads to biased estimation. Another example comes from the fact that previous studies demonstrate that richer human capital predicts a higher probability of using contacts to find work (e.g. Bentolia, Michelacci and Suarez 2010; Lin and Ao 2008). If human capital is not controlled for, it actually enters into the error terms of both the substantive model and the selection model, thereby giving rise to biased estimation. As Marsden and Hurlbert (1988) stress in estimating the causal effect of social capital in the labour market, there are substantive reasons to think that unmeasured factors predisposing people to use contacts may be correlated with error terms in the job outcome equations.

To get around the sample-selection problem, an important work implementing Heckman's two-stage method was conducted by Marsden and Hurlbert (1988). Drawing on Berk (1983) and Berk and Ray (1982), they use a logit model to predict the probability of using contacts on a constellation of observed variables and then include the predicted probability of using contacts into the substantive model as a control.[15] Following them, Wegener (1991) also uses this method and controls for the exclusion probability of not using contacts and not changing jobs. More recently, Lin (2003) uses both the probit and logit models to predict the

15 According to Berk (1983:394), the hazard rate calculated from the predicted probability is treated as a new variable and included in any substantive equations. However, the hazard rate from the logit model is simply the predicted probability.

propensity for using contacts and then controls for it so as to alleviate the sample-selection bias. The study confirms the causal effects of used social capital in the Chinese labour market.

Furthermore, as Wooldridge (3rd ed. 2006:620) mentions, to practically distinguish the sample selection from a mis-specified model, the constellation of explanatory factors in the substantive model should be a strict subset of the constellation of explanatory factors in the selection model. That is, any independent variables appearing in the substantive model should be included in the selection equation. Moreover, one should have at least one explanatory variable in the selection model that has no partial effect on job outcomes. Although this is not absolutely necessary in order to apply the Heckman methods, without imposing such exclusion restrictions it will be extremely difficult to distinguish the sample selection from a mis-specified function from the substantive model. For example, in Lin (2003) the variables appearing in the selection model but not in job outcome model are work experience and the rank of father's work unit education. In Marsden and Hurlbert (1988), the exclusion restriction is job experience. However, one may have good reason to suspect that these variables (e.g. work experience) should also appear in the substantive model.[16] In other words, the exclusion restrictions are not of good quality so that there may still be estimation bias in these applications.

16 For instance, the Mincer equation in labour economics includes experience as a major explanatory variable to predict wages.

2.3 Estimating the Effects of Accessed Social Capital

Durlauf (2002) examines how to identify the causal effects of accessed social capital under two different assumptions: one is that accessed social capital is predetermined as a group characteristic and the other is that accessed social capital is measured as the group behaviour/outcome and contemporary to the individual behaviour/outcome as dependent variables. As he points out, the social capital model under the former assumption is the same to identify the exogenous effect and the latter to identify the endogenous effect.[17] Empirical studies referred to this chapter can thus be classified into two strands in light of Durlauf's dichotomy (Table 3 and Table 4 respectively summarize the selected empirical studies along these two strands).

2.3.1 The Exogenous Effects of Accessed Social Capital

This strand of literature assumes no presence of social multiplier, thereby focussing on social capital's exogenous effects. That is, the measure of social capital is not the group-level measure (e.g. mean or sum value) of the individual

17 Note that using the terminology of exogenous effects does not mean that the social capital variable is assumed to be exogenous, but is rather used to distinguish it from the endogenous effects (Manski 1993). To understand this, one could always use the terminology of contextual effects to replace the exogenous effects. This study uses the terminology following Manski (1993).

job outcome as the dependent variable. As Table 3 shows, three types of social capital variable have been used. The first one is a direct measure of embedded resources influencing job opportunities (Panel 1). Accessed social capital is thus often operationalized as the group indicator of social resources embedded in the whole social network (e.g. job status, wages or educational attainments). The second type of social capital variable is a structural measure such as network size, network diversity and network constraints in terms of the respondent's social network (Panel 2). Finally, the third type combines the above two, jointly using these measures to predict job outcomes (Panel 3).

As can be seen in Panel 1, Lin and Dumin (1986) and Völker and Flap (1997) use the resource reachability (the highest occupational prestige score of contacts in the social network) as a measure for accessed social capital, and report a positive influence on individual occupational prestige. Flap and Völker (2001) examine how the goal-specific social capital – a network with a given structure and content – affects job satisfaction. They find that accessed social capital is not an all-purpose good within the workplace. In particular, networks of strategic, work-related ties increase the employee's job satisfaction in relation to "instrumental aspects" (e.g. income), while networks of identity-based solidarity ties improve job satisfaction in terms of "social aspects" (e.g. social climate at work and cooperation). Sørensen (2000) uses the data from commercial television stations in 24 broadcast markets between 1953 and 1988 to test

whether a history of group demographic heterogeneity influences individual turnover rate. Using the over time mean of standard deviation of manager team tenure as a predictor, he fits a piecewise-constant hazard rate model and reveals that those belonging to heterogeneous groups throughout their tenure are more likely to exit the team.[18] Based on two different cross-sectional data sets, Mouw (2003) also finds that the average educational attainment and the proportion of friends with steady jobs have positive effects on the respondents' wages. Yet he aptly argues that, with the possible presence of both self-selection and omitted variable bias, we cannot determine whether these findings reflect the causal effect of accessed social capital or the fact that similar individuals tend to associate with each other.

To deal with the endogeneity problem that might plague estimations, approaches such as the fixed effects model, modelling selection process and IV have been used in recent studies identifying the exogenous effects of accessed social capital in the labour market. For instance, Weinberg, Reagan and Yankow (2004) explore the role of a set of neighbourhood characteristics and job proximity on annual work hours. They use the exact street addresses recorded annually for respondents to the NLSY 1979 to match the neighbourhood social characteristics of the relevant respondents from the 1990 Census of Population. Exploiting the constructed panel data, they estimate a fixed effects model. Holding individual

18 This is also called 'Event History Analysis'.

heterogeneity as fixed, the relationship between neighbourhood employment rates and individual employment is reduced by two-thirds in relation to the corresponding OLS estimates.

Table 3: Estimating the Exogenous Effects of Accessed Social Capital

Researches	Independent	Dependent	Data & Model	Finding
	1. Direct Resources Measure			
Lin & Dumin 1986, (USA)	Accessed highest status; Range	Status	Cross-section OLS	Positive
Völker & Flap 1999 (GDR)	Accessed highest status	Status	Cross-section OLS	Positive
Flap &Völker 2001 (Holland)	Goal-specific social capital	Job satisfaction	Cross-section OLS	Positive
Sørensen 2000 (USA)	Mean of the SD of group tenure (from entry year to t-1)	Individual turnover	Panel piecewise-constant hazard rate model	Positive
Mouw 2003 (USA)	Average education; Proportion of friends with steady job	Wages	Cross-section OLS	Positive
Weinberg, Reagan & Yankow 2004 (USA)	Neighbourhood characteristics Neighbourhood job proximity	Annual work hours	Panel Fixed Effects	Positive
Bayer & Ross 2009 (USA)	Geographic proximity to job Neighbourhood poverty rates Neighbourhood education	Employment Weeks worked Weekly hours	Cross-Section OLS+IV	Positive
Dujardin & Goffette-Nagot 2007 (France)	Location in a deprived neighbourhood	Unemployment rate	Cross-section Modelling selection	Positive
	2. Network Structural Measure			
Burt 1992 (USA)	Structural holes	Promotion rate	Cross-section OLS	Positive

Podolny & Baron 1997 (USA)	Size network	Intra-organization promotion	Cross-section OLS (Heckman correction)	Positive
Erickson 1996 (Canada)	Diversity in accessed positions	Job autonomy	Cross-section OLS	
Kim 2001 (USA)	Network constraint (density) Network bridge (diversity) Number of ties with elite school background; Number of ties to notables	Annual earning	Cross-section OLS	None None Positive Positive
Korpi 2001 (Sweden)	Network size of unemployed	Probability of Employment	Cross-section Cox Model (Heckman correction)	Positive
Munshi 2003 (Mexico)	Network size in community	Employment Wage	Panel Fixed effects IV	Positive
Yakubovich 2005 (Russia)	Tie strength	Employment	Cross-section Fixed effects (within-subject)	Positive
Beaman 2008 (USA)	Network size Number of long-tenure contacts (current or prior)	Employment Wage	Panel Fixed-effects	Negative Positive
Hellerstein, McInerney & Neumark 2009 (USA)	Network connectedness	Working at the same business establishment	Cross-section Matched pairs	Positive
	3. Resource + Network			
Campbell, Marsden & Hurlbert 1986 (USA)	Factor score of diversity, size and density. Mean/max education and prestige	Status Family income	Cross-section	Positive
Flap & Boxman 2001 (Holland)	Sum of the product of tie strength and resources; Sum of the product of person's autonomy and resources; Network structure	Income	Cross-section OLS	Positive
Oreopoulos 2002 (Canada)	Neighbourhoods size; Construction style; Neighbourhoods poverty rate	Education attainment Earning; Income	Natural Experiment	None
Seibert, Kraimer, &	Weak ties; Structural holes;	Career success	Structural	Positive

Liden 2001 (USA)	Access to information; Access to resources; Career sponsorship.		equation	
Ruiter and De Graaf 2009 (Holland)	Status of the involved voluntary associations; Diversity of voluntary associations	Wages and Status	Life-course data Multilevel analysis	Positive

Notes: Positive= Positive effects and significant at 0.05 level
Negative= Negative effects and significant at 0.05 level
None= Insignificant effects
SD= Standard deviation

Using data taken from the 1999 French Census, Dujardin and Gpffette-Nagot (2006) examine how individual unemployment is affected by living in a deprived neighbourhood. To get round the endogenous group member problem plaguing causal inference, they estimate simultaneously one probit model for unemployment and another probit model for the choice of neighbourhood (deprived or not deprived).[19] Thus, the simultaneous probit model accounts for the unobserved factors influencing both unemployment and residential choices by explicitly estimating the correlation matrix of residuals. They find that, although the neighbourhood effects are mitigated when dealing with sorting to neighbourhood, living in the 35% more deprived neighbourhoods significantly increases the probability of unemployment.

Bayer and Ross (2009) investigate exogenous neighbourhood effects on labour market outcomes such as labour market participation and weeks worked per

19 The public housing accommodation is included in the selection model, since it is argued to be a strong determinant of location in deprived neighbourhoods in France. Since getting public housing is not random, Dujardin and Gpffette-Nagot (2006) in fact estimate a third probit model for housing tenure and use the household's observed demographic characteristics as the exclusion restrictions.

year. First of all, they exploit the monotonic relationship between neighbourhood housing prices and neighbourhood quality in equilibrium. In doing so, they use a certain function of neighbourhood housing prices as a control for the neighbourhood unobservables. This therefore alleviates the omitted variable bias at the group level and the self-selection problem. Second, to deal with individual-level unobservables, they instrumented for each individual's observed neighbourhood characteristics with the mean neighbourhood characteristics for a set of observationally identical individuals. The justification for this is that the proposed IVs are related to location because similar individuals make similar neighbourhood choices, and they are not correlated to individual's unobservable factors because the IVs have been constructed from observable characteristics assumed to be uncorrelated with individual unobservables.[20] Using confidential microdata from the 1990 Decennial Census for Boston, they show that the direct effects of geographic proximity to jobs, neighbourhood poverty rates, and average neighbourhood education are substantially larger than the conditional correlations identified using OLS, although the net effect of neighbourhood quality on labour market outcomes remains small.

We turn now to Panel 2, where the second strand of literature primarily focusses on how the structure of social networks, as accessed social capital, influence job outcomes. Ground-breaking studies such as that by Burt (1992) demonstrate the

20 It is a standard assumption in a single equation model that all the explanatory variables should be exogenous.

power of network structure (e.g. structural holes indicating the absence of connections among those in the network). Other early research focusses on the effects of network scale, diversity, density and many other network indicators reflecting access to job-related information and influence. Most findings espouse a positive association between the accessed social capital and various job outcomes. When it comes to the estimation strategies used to overcome the endogeneity problem, a constellation of studies are worth mentioning here. For example, Podolny and Baron (1997) consider the possible sample-selection bias in responding to surveys and implemented Heckman methods in recompense. Based on the data from a high-technology engineering and manufacturing corporation, they then report that the promotion (grade advancement) within an organization is enhanced by having a large, sparse network of informal ties. Using the similar estimator suggested by Heckman and Robb (1985), Korpi (2001) deals with both the possible self-selection and sample-selection problem when exploring the effects of network size on job search among the unemployed in Sweden. The reason of doing so is two-fold: first, unemployed persons are likely to consider the net benefits of talking about job search with certain particular types of friend and, second, samples with the transition experience from unemployment to employment may entail sample selection. He then reports that the network size is positively correlated with the probability of labour market participation.

The fixed effects model and matched-pair method are also used to estimate the role of structural measures of accessed social capital. For example, using the within-subject method with several observations per person, Yakubovich (2005) collects data on multiple potential contacts for each worker and substantiates the positive effects tie strength has on individual employment in Russia. Hellerstein et al (2009:5) use matched employer-employee data to see how network connectedness influences the probability of working in the same business establishment. The measure of social networks is operationalized as "the extent to which employees of a business establishment come disproportionately from the same sets of residential neighbourhoods (defined as census tracts), relative to the residential locations of other employees working in the same census tract but in different establishments". Based on the 2000 Decennial Employer-Employee Database, the findings show that the network connections among workers in the residential location positively predict the possibility that workers are employed in the same business.

Two studies combining more than one correcting method in order to deal with the endogeneity problem offer good templates for casual inference in social capital studies. Munshi (2003) includes individual fixed effects and uses rainfall in the community of origin to instrument the size of the immigrant's network in the US. He then successfully verifies the positive impact of immigrant's network size on employment probability and wages in nonagricultural jobs. The IV

estimate is larger than that established through OLS. Munshi (2003) accounts for this by arguing that there might be established migrants who finally move back to Mexico. Beaman (2008) studies the dynamic implications of social networks for the labour market outcomes of political refugees resettled in the U.S. To circumvent the endogeneity problem, he exploits the fact that the refugees are distributed across cities by the resettlement agencies in a random way. That is, this is an experimental application. In addition, the variation in social network size across cities and ethnic groups over time facilitates the fixed effects strategy, ruling out unobservable factors at the city and nationality group levels. Combining these two strategies, he then demonstrates that the number of social network members resettled in the same year (or one year prior) is negatively related to labour market outcomes (i.e. employment and hourly wage). However, the network size of long-tenured peer refugees positively predicts job outcomes.

Panel 3 in Table 3 demonstrates the third strand of literature probing the effects of both direct resource measures and structural measures at the same time (e.g. Campbell, Marsden and Hurlbert 1986; Flap and Boxman 2001). In the early studies, limited attention was paid to the endogeneity problem. It is worth mentioning the use of the structural equation model (SEM) to trace the relationship among a constellation of social capital variables. Seibert, Kraimer and Liden (2001) conceptualize social capital as both network structure (e.g. network size, structural holes and weak ties) and social resources measured by

access to information, access to resources, and career sponsorship. The results of SEM confirm a positive effect from social capital on job success regarding promotions, salary, and career satisfaction. However, despite SEMs being useful in dealing with measurement error and latent variables with multiple indicators, they encourage confirmatory rather than exploratory modelling. As a result, SEMs cannot address the endogeneity problem without "making assumptions about the underlying causal structure, assumptions that are ultimately untestable and subject to second guessing" (Mouw 2006:93).

A study based on a natural experiment can be expected to alleviate the endogeneity problem. Oreopoulos (2003) exploits the fact that certain adults born between 1963 and 1970 and assigned as children to reside in a housing project in the Toronto metropolitan area were randomly assigned to the neighbourhood. Comparing indicators of adult socioeconomic status (e.g. earned income, likelihood of unemployment and welfare participation) among neighbourhoods with different sizes, construction styles (high-rise vs. low-rise) and poverty rates, he finds no significant differences and the estimated correlation in outcomes for those who lived in the same housing project but within different families is less than 0.005. Therefore, when residential locations are not selected, neighbourhood characteristics (including both direct resource measures and structural measures) have a tiny role.

Based on life-course data sets in the Netherlands, Ruiter and De Graaf (2009) report that the characteristic of the involved voluntary association pays off, controlling for previous socioeconomic characteristics at the time of a job start to alleviate the potential endogeneity problem. In addition, focussing on the attributes of social network generated by voluntary association involvement entails they better avoid the selection problem, given that the status of a person in terms of voluntary associations is generally not directly related to that person's job.

2.3.2 The Endogenous Effects of Accessed Social Capital

Two types of endogenous effects studies, namely regressing on nonlinear models (Panel 1) and on linear-in-means models (Panel 2) are both presented in Table 4. Note that the dependent variables of interest are not continuous variables but dummy variables reflecting individual work choices in Panel 1, while Panel 2 is otherwise. According to Brock and Durlauf (2001), the non-linearity in the logit or probit models can be used to overcome the reflection problem. In particular, "the binary choice framework imposes a nonlinear relationship between group characteristics and group behaviours whereas the linear-in-means model (by definition, of course) does not" (Brock and Durlauf 2001:254). Thus, the potential sources of endogeneity for the studies presented in Panel 1 are primarily conventional omitted variable bias, simultaneity bias, and selection

bias. In contrast, for studies in Panel 2, the reflection problem should also be a concern.

Case and Katz (1991) consider that the neighbourhood family background as an exogenous factor should also be included in the model where neighbourhood outcome is an explanatory variable. By controlling for a set of predetermined observed family and neighbourhood factors, they report a significant neighbourhood peer effect on the propensity of youths to be out of work. Nevertheless, one rather suspects that there may be unobserved neighbourhood-level, household-level and individual-level factors. Building on these earlier studies, more sophisticated methods have been used to correct the potential bias in recent research .

Based on unique Swedish panel data, Hedström, Åberg and Kolm (2003) estimate a Cox proportional hazards model to explore how the hazard of individuals leaving unemployment is influenced by the unemployment level in neighbourhoods. To control for time-invariant unobservable characteristics of the neighbourhoods and sorting to residence location, they estimated a fixed-effect specification including a constellation of observed factors and nearly 700 dummy variables – one for each neighbourhood (except one). As they point out, despite the risk that this way of controlling for between-neighbourhood differences may downwardly bias the true neighbourhood effect, the findings

even with these "excessive controls" show that a substantial social-interaction effect exists and indeed confirm the role of peer influence.

Fixed effects method aside, the IV approach is employed as well to get around the potential omitted variable bias, selection bias and simultaneity bias. Contreras et al. (2007) use the IV approach to explore the neighbourhood effects in determining the participation of Bolivian women in income-generating activities. The estimation results suggest that neighbourhood peer effects can be a mechanism through which women obtain access to salaried jobs in Bolivia. Maurin and Moschion (2009) examine how neighbours' labour market participation affect a mother's entry decision in France. To alleviate the endogeneity problem, they use the average children sex-mix of neighbours to instrument the neighbours' labour market participation. The justification of the IV lies in the fact that, in France, the children sex-mix has a significant impact on a mother's decision to participate in the labour market but does not affect the neighbourhood choice. Their findings demonstrate that there is significant elasticity of own labour market participation in relation to neighbours' participation.

In addition, rather than directly measuring the magnitude of the endogenous effects of accessed social capital, there are also some innovative methods to confirm the existence of endogenous effects. For example, Araujo, Janvry and

Sadoulet (2010) analyse the role of peer effects in helping rural households obtain access to off-farm non-agricultural employment. Based on Mexican data, they use statistics from a set of models to reject the possibility that village peer effects measured by mean village participation rate can be proxies for unobserved individual, village-level, or individual type effects. That is, although they have not directly measured the magnitude of the role of accessed social capital, they confirmed that social capital's endogenous effects in the labour market do exist.

Table 4: Estimating the Endogenous Effects of Accessed Social Capital

Researches	Independent	Dependent	Data and Model	Finding
	1. Non-Linear Model			
Case and Katz 1991 (USA)	Neighbourhood Idle rate (no labour-force participation)	Idle status	Cross-section	positive
Hedström, Åberg & Kolm 2003 (Sweden)	Mean employment rate	Employment	Panel data Fixed effects Cox model	Positive
Araujo, Janvry & Sadoulet 2010 (Mexico)	Mean village participation in off-farm non-agricultural work	Participation	Cross-section Rejection Method	Positive
Contreras et al. 2007 (Bolivia)	Mean employment rate; Mean waged-work rate	Employment Waged work	Cross-section IV + Proxying	Positive (female)
Maurin & Moschion 2009 (France)	Mean labour-market participation of the other mothers in the neighbourhood	Mother's LM participation	Cross-section IV	
	2. Linear-in-Means Model			
Marmaros & Sacerdote 2002	Mean wage	Wages	Cross-section Random assignment	Positive

(USA)				
Laschever 2005 (USA)	Mean wage Mean unemployment rate	Wage Unemployment	Cross-section Multiple Reference Group (see below)	Positive
Shvydko 2007 (USA)	Mean earnings of firm peer	Individual earning	Panel Fixed effects	Positive
Grodner & Kniesner 2007 (USA)	Mean group hours worked	Hours worked	Cross-section IV	Positive

Notes: Positive= Positive effects and significant at 0.05 level
Negative= Negative effects and significant at 0.05 level
None= Insignificant effects

As for the linear-in-mean models in Panel 2, where a simultaneity bias as well as a reflection problem arise as the leading estimation challenges, some studies just estimate the mixed effects (combining the exogenous and endogenous effects) rather than separating them from each other. In other words, they deal with the simultaneity bias but leave the reflection problem still open. For example, Marmaros and Sacerdote (2002) exploit random assigning to accommodation in Dartmouth College and find a large positive association between getting help from fraternity/sorority contacts and the prestige of the obtained job. However, as they concede, due to what we have seen Manski (1993) call the reflection problem, they cannot interpret the estimated coefficient directly as the actual size of the endogenous peer effect. Likewise, to estimate the impact of mean co-worker earnings on the individual's earnings, Shvydko (2007) fits a fixed-effects model on panel data to alleviate the omitted variable bias and self-selection problem. Nevertheless, he cannot separate the endogenous and exogenous effects because the simultaneity problem has not been addressed.

Yet there are some studies that successfully distinguish between the endogenous effects and the exogenous effects via innovative methodologies. For example, Laschever (2005) analyses the employment effect of social networks of WWI veterans by introducing a new framework named Multiple Reference Group (MRG) to help identify the model. The attraction of MRG lies in that it allows for the fact that some people belong to more than one social network. Consider there is a subpopulation AB belonging to two groups A and B (i.e. AB is the overlapping section of A and B), and the rest of the people in A and B are called RA and RB respectively. Importantly, individuals in RA do not know anyone in RB, but RA and RB do share a common intermediate group of contacts, namely AB. As a result, one would not expect the characteristics of RA to directly influence the outcomes of RB since they are not directly tied. If, therefore, one can find an effect, it must be evidence for the existence of an endogenous effect that the intermediaries AB had propagated. That is, suppose the effect of RA on RB is captured by E1, and the effect of RA on AB is captured by E2, the ratio of E1/E2 finally leads to identification of the endogenous effects. Using this novel method, Laschever exploits the fact that the veterans in the sample belong to both a military unit (in which he served during WWI) and the neighbourhood (the group of people he was geographically closest to in the 1930 census) as two reference groups. He then identifies the endogenous effect and finds it to be much stronger than the exogenous/contextual effects.

With a similar strategy, Grodner and Kniesner (2007) examine the social interaction effects that the average hours worked by the reference group have on the workers' individual hours worked. Since the reference group in their study contains those individuals in "close economic distance", which is measured by factor analysis, each respondent has an individual-specific reference group. Thus, they exploit the fact that there are workers who belong to more than one reference group so that hours worked by people in the adjacent reference group can be used as an instrument. Specifically, the IV is measured by a social grid with two social coordinates from factor analysis. The justification for using it as the IV is that it is correlated with the mean hours worked in the individual's reference group, since individuals in the group and the adjacent reference group belong to the same economic neighbourhood. Meanwhile, the IV should be uncorrelated with individual unobservables because the particular individual does not belong to the adjacent reference group.

3. CONCLUSION

Over the past three decades, social scientists have recognized the importance of social capital in a variety of contexts, such as crime, labour markets, education, welfare use, substance use, social learning and adoption of innovations. One common feature in the relevant studies is the argument that individual behaviour and outcomes are influenced by the others in a given social network via different

mechanisms (e.g. information sharing, normative pressure or imitation). Put simply, social capital matters. In the specific context of labour markets, a vast body of literature has attempted to provide statistical estimates of the magnitude of social capital's effects on individual job outcomes. Although sociologists have argued, long before economists, that social capital matters, systematically dealing with the endogeneity problem which is pivotal to causal inference has long been ignored in sociological literature. This has greatly undermined the explanatory power of social capital theory. In contrast, as Angrist and Krueger (1999) argue, labour economics has come to be distinguished by the use of cutting-edge methods for identification. Only recently has some sociological literature focussing on social capital called for attention to the model-identification problem (e.g. Mouw 2003, 2006).

Having arrived at this point, it is feasible and appropriate to reflect on what has so far been accomplished in terms of estimating the causal effects of social capital in the labour market. As the debate on whether social capital really matters for job outcomes has not yet been settled, this chapter also surveys leading empirical studies in econometrics on the basis of their contribution to the identification of strategies to address the endogeneity problem. The discussion shows that the endogeneity problem is pervasive yet not unsolvable. In this literature review, the major sources of endogeneity – omitted variable bias, self-selection bias, sample-selection bias and simultaneity bias – are all addressed. The major

identification strategies (e.g. the IV approach, matched pairs, the fixed-effects model, the random assignment to treatment and natural experiment) employed to mitigate the endogeneity bias in social capital studies are discussed with respect to their ability to credibly improve the estimation of the role of social capital in the labour market.

Considering the pervasive presence of the endogeneity problem, however, can we now arrive at a conclusion that the estimated effects of social capital in previous sociological studies are in general overstated or understated? I think it is still premature to make such a conclusion. Theoretically, to establish how the estimates are biased, one should fully examine how the unobservables are correlated with the individual outcomes of interest and the social capital variable. Taking accessed social capital for example, when there is self-selection driven by social homophily in the formation of the social network, the estimated effect of accessed social capital would be upward-biased. This is because the estimates must represent a combination of the "real" role of network resources and the effect of the unobservable on which social homophily arises. In contrast, if self-selection is driven by social heterophily, things are very different. Suppose that individuals of poor ability are aware of their disadvantages in the labour market. As rational actors, these individuals are therefore likely to select friends possessing more social resources. In this case, a standard estimate of the role of accessed social capital will be downward biased, as it mistakenly offsets the role

of the unobservables. This type of heterophlious selection is not impossible, indeed being supported by some recent empirical literature (Bayer and Ross 2009).

So, whether social capital is operationalized as using contacts, used social capital or accessed social capital, empirical applications should explicitly consider how to deal with the endogeneity problem in the context of their data. Drawing upon the existing studies and listing all the identification challenges and strategies, the intention of this review is neither to pick on earlier sociological analysis nor to discourage researchers. Rather, it aims to call for more attention to this discussion and for further research to be carried out. Exactly as Mouw (2006:100) points out, "although the papers reviewed here are dispersed across several disciplines and substantive topics, their creative methodological approaches may provide a useful template for future researchers to arrive at a more definitive conclusion regarding the effect of social capital on individual outcomes". Notice that with observational data, it is almost impossible to fully address the endogeneity problem. However, sociological analysis of social capital should take the endogeneity problem very seriously. Even without informative data sets, one should at least tell one's readers that the empirical findings are threatened by potential bias led by a correlated component of the error term in the model. Given informative data sets, one should carefully compare the results obtained from standard single-equation models and those from advanced models which

are capable of dealing with potential endogeneity problem. This helps to understand the underlying social mechanism through which social capital affects the individual outcomes of interest.

Finally, a related issue is how to choose appropriate identification strategies. There should be no doubt that panel data provide more information from which we can exploit the time variation of the explanatory variables and the outcome variable to identify the causal effects. This is especially so when estimating the role of using contacts or used social capital. The reason is straightforward: people changed jobs, so they used contacts and mobilized social capital more than once. It is therefore possible for sociologists to collect retrospective information and get panel data. In Chapter 5 and Chapter 6, I compare the estimates from the OLS models and the estimates from first-difference models. The panel analysis is possible because the data sets that I work on are of two time-periods for each respondents.

However, sociologists focussing on network social capital are often either provided with cross-sectional data, or hard-to-conduct within-analysis. There are two major reasons for this. First, the members and structure of the ego-centric social networks can be very stable through a long time. Consequently it is often impossible to get panel data on social networks. Second, within-group analysis is useless if there was only one observation for each network/group in a given

dataset. Chapter 3 and Chapter 4 are cases in point: the data are available merely on each "ego" of the ego-centric networks. As a result, what is needed is an IV that influences the dependent variable only by affecting the social capital variable. As can be seen in the following chapters, the IV approach is used in chapters 3 and 4 to correct for endogeneity bias.

As we have seen, identification strategies can be combined to better alleviate the estimation bias. In this thesis, I combine the first-difference method with the Heckman correction in Chapter 6 in order to deal with the additional self-selection problem. I also combine the Heckman method with the IV estimator in Chapter 3. Notice that one should always explicitly illustrate the relevant assumptions needed for estimating advanced models. For instance, when using the IV method, one should clearly establish the validity of the IV and convince the readers regarding the assumption of zero correlation with the error term. When fitting fixed effects models, one must bear in mind that the analysis is based on the untestable assumption that all relevant unobserved inter-individual differences are time constant. In short, we should not make assumptions unblinkingly. Furthermore, since these assumptions are often untestable, it would be very useful to compare those estimates generated from standard single-equation models with those from advanced models. This is exactly the reason why OLS and probit estimates are also presented in the empirical chapters in this thesis.

Chapter 3

Network Effects among Rural-to-Urban Migrants in China

1. INTRODUCTION

Social networks are important for the analysis of migrants because they can be expected to influence migration decision-making as well as its outcomes. Although the role of social networks in migration has been widely studied (e.g. Uhlenberg 1973; Ritchey 1976; Banerjee 1983, 1984; Taylor 1986; Massey 1987; Boyd 1989; Fawcett 1989; Carrington, Detragiache, and Vishwanath 1996; Gurak and Caces 1992; Massey et al. 1993; Palloni et al 2001; Bauer et al. 2002; Munshi 2003; Curran and Rivero-Fuentes 2003; Haug 2008;), sociological analyses of how networks influence internal migration under state socialism are very scarce. In contemporary China, labour migration from rural to urban areas influences the lives of hundreds of millions of people. A conservative estimate from the National Bureau of Statistics of China (NBSC) (2008) indicates that the total rural-to-urban migration flow had reached 0.14 billion in 2007.

Like international migrants, the social networks of rural-to-urban migrants in

China are mainly based on kinship and common village/township (see Massey et al. 1993 for international migration; see Rozelle et al. 1997, Zhang 2001 and Zhao 2003 for internal migration in China). This chapter aims to assess whether a migrant's wage at the destination is positively influenced by the size of his or her village-based migrant network. The core hypothesis behind this study is that the total number of fellow villagers who migrated to urban areas positively influences one's wage. The underlying mechanism is straightforward: a wider origin-based network gives an individual more job-related information and hence better wage. Notice that the size of network used is not measured in terms of administrative boundaries at the destination, but on the origin villages. This is mainly because previous findings (Zhao 1999, 2001; Hare 1999) reveal that rural-to-urban migrants in China move frequently between origins and destinations, strongly implying the pivotal role of the village-level outflow of migrants in facilitating job-match by transmitting useful job leads.

Although it is intuitive to expect that migrant networks will bring about benefits, identifying causal effects is not easy with observational data. To systematically deal with potential estimation biases, this chapter estimates and compares three models, namely an ordinary least squares (OLS) model, a Heckman's two-stage model and an IV-Heckman model using natural disaster as the instrument. To the best of my knowledge, this is the first application of instrumental variable (IV) technology in sociological analyses seeking to estimate network effects. The

result of the empirical analysis demonstrates that the size of the origin-based network significantly improves the wages of its member migrants.

The rest of the paper is structured as follows: Section 2 conducts a brief review of the literature on labour migration in China and provides the theoretical background to the analyses. Section 3 addresses the models and pays close attention to the identification strategies used to correct for endogeneity biases. Section 4 describes the data and definitions of relevant variables. Section 5 presents estimation results and Section 6 proposes interpretations. Finally, Section 7 concludes.

2. BACKGROUND

Internal migration has changed China in pervasive and profound ways, given the vast number of people moving from undeveloped rural areas to cities. The magnitude and rapid growth of internal migration from 2000 to 2007 is illustrated in Figure 1[21]. For panoramic surveys on rural-to-urban labour migration in contemporary China, interested readers can refer to the migration worker report of the Research Office of State Council (ROSC 2007), Li (2008), De Brauw and Giles (2008), and Liang (2006).

21 According to the NBSC, rural-to-urban migration is defined as working outside of one's township for more than one month in a year. Given the different definitions of migrant population used in surveys and censuses, estimations of China's migrant population vary. However, the huge population and rapid growth is a stylized fact.

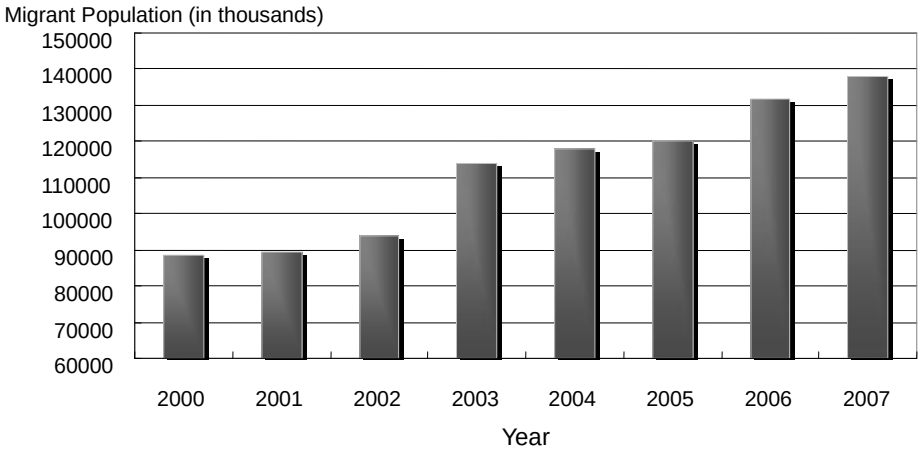


Figure 1. Rural-to-urban migration flow in China by year, 2000 to 2007

2.1 Origin-based Networks and Job Searching

Whether the backdrop is a market economy or a transitional economy under state socialism, it has long been recognized that many persons find jobs through their contacts or relatives (Granovetter 1974; Montgomery 1992; Bian 1997). Numerous social capital and network effect studies reveal that embedded resources lubricate the job matching process and affect job outcomes (see Lin 1999 for a review) by offering job-related information and influences. As for migrants, their social networks at the destination are mainly "sets of interpersonal ties that connect migrants, former migrants, and non-migrants in origin and destination areas through ties of kinship, friendship, and shared community origin" (Massey et al. 1993:448). These networks provide job-search assistance for migrants by relaying useful job leads and reducing searching

costs[22] (Banerjee 1983, 1984; Taylor 1986; Massey 1987; Massey et al. 1993 etc.).

This is also the case in internal migration in China. Rural-to-urban migrants in China in particular have limited ties with the urban communities and little access to institutional support at the destination, so they have to heavily rely on their origin-based networks to find jobs (e.g. Rozelle et al. 1997; Solinger 1999; Zhao 2003). For instance, surveys show that 70 per cent of rural-to-urban migrants found jobs though the network of village-based friends or relatives (Meng 2000). Importantly, previous migration studies in America and Europe have revealed that more job-related information from origin-based network gives a migrant more choices for jobs, and hence, enables the individual to find a higher earning position relative to others with narrower networks (Massey and Espinosa 1997; Munshi 2003; Rainer and Siedler 2008). It is therefore of interest to explore whether this pattern holds for internal migration in China.

2.2 Village Outflow as the Origin-based Migration Network

According to Hare (1999), labour migrants in China tend to have little attachment to destination cities and return home frequently during a year. An average migrant is reported to return home two to three times annually. For example,

22 There is also evidence showing that, even though migrant networks help to find temporary jobs, the inability of rural migrants to use the networks to share resources is ultimately detrimental for participants' long-term outcomes (e.g. Balderrama and Molina 2009).

most of the migrants go home to be reunited with their family members during the Lunar New Year. During the busy seasons, such as seedtime or harvest time, many migrants go back home to help household members to raise or harvest crops. Surveys also show that the majority of migrant workers spend less than nine months in their host cities (Zhao 1999). In addition, permanent returnees are very few in China because migrants tend to move out again after returning (Hare 1999; Zhao 2002, 2003). Indeed, rural-to-urban migration in China is actually defined by its "floating" nature (see e.g. Chan and Zhang 1999)[23]. In general, most Chinese rural households regard migration as a short-term arrangement (Zhao 1999). Since rural-to-urban migrants in China return home frequently in a year, villages of origin serve as an important intermediary for job leads to be transmitted among villagers.

It is natural to expect there to be an important role played by origin-based networks at the destination, a subsection of the total village outflow. For example, fellow villagers who migrated to the same city can help each other in the job-seeking process. However, since previous findings suggest that rural-to-urban migrants at different destinations return home frequently even in a year, the total migration outflow at the village level can be suspected to

23 Why do migrants choose not to permanently settle down in a given city given the tremendous rural-urban income gap? The main reason is the presence of the Household Registration System inherited from the planned-economy era. Despite migrants being *de facto* urban residents, there are very few channels by which migrants could transfer from "agricultural" to "non-agricultural" residential status (Wu and Treiman 2007). In addition, since farmland cannot be sold and the process for renting is quite complicated, peasants are reluctant to give it up (Roberts 2000).

constitute a larger origin-based network, influencing the labour market outcomes of its members. It is therefore of special interest to assess the impact that the outflow of migrants at the village level has on the wages of migrants at the destination.

2.3 Migration as a Selection Process

The prime determinants of migration decisions can be classified into two types, namely push and pull factors. In particular, the expected wage gap between sending and hosting areas is the major pull factor, while surplus rural labour is often viewed as the major push factor (see, for example, Todaro 1969; Harris and Todaro 1970; Robinson and Tomes 1982). Following this line of research, a large body of China studies focus on the network effects on migration decision-making (e.g. Hare and Zhao 2000; Mallee 2000; Meng 2000; Zhao 1999, 2003; Bao et al. 2007; Chen et al. 2008).

Rural-to-urban migration in China is first of all driven by the income gap between rural and urban areas. This income difference after the reform and opening-up narrowed in the early 1980s but has widened again since 1985 onwards (Zhu 2002, Zhao 2003). The surplus labour in rural China is argued to be one of the push factors (Zhao 1999; Chen et. al 2008). In addition, a sizable literature on the determinants of internal migration in China finds that a set of individual

characteristics (e.g. health, age, gender), household and village level attributes (e.g. household labour force, shortage of farmland, local rural taxation policy, distances to cities) matter for migration as well (see e.g. Cai 1996, Hare 1999 and Zhao 2003;).

Importantly, these studies strongly suggest that migration is a self-selection process based on a set of local, household and individual factors. Given the huge income-gap between rural and urban China, it is natural to suspect that persons who are more "able" in terms of unobserved earning ability are more likely to migrate. In fact, previous China studies (Zhao 1999, 2003) show that observed individual characteristics signalling higher earning ability (e.g. gender, education, and health) positively predict migration. Considering this, the selection bias should be taken seriously when analysing the causal effects of migrant networks.

3. MODELS AND IDENTIFICATION ISSUES

The specification of the model should include the size of the network as an explanatory variable and observed individual, village and destination characteristics as controls. However, naïvely fitting an OLS model to estimate the network effects is problematic due to the presence of endogeneity.

3.1 Sources of Endogeneity

To consistently estimate an OLS model, the network size variable should be uncorrelated with unobservables. This assumption, however, is very likely to be violated due to there being various sources of endogeneity in the context of this study. There are four main concerns.

First, as mentioned before, there might be a sample-selection based on unobserved abilities/skills in migration decision-making. The decision to undertake migration is determined by a set of individual-, household-, original- and destination-level attributes. If a certain unobserved factor influences both the migration decision and the job outcomes, the estimation will be biased due to there being an unrepresentative sample.

Second, the wages of migrants can be region-specific. If the destination city dummies are not included in the model, the estimation may be biased because favourable labour market conditions at the destination may not only result in a higher wage for a migrant but also attract more fellow villagers from his or her origin village. As will be shown in the next section, the data used does not contain any information regarding destinations. This strongly implies the presence of an omitted variable bias using the naive regression model.

Third, unobserved village-level factors can also bias the estimates by influencing

both the wage of migrants and village migration outflow. In empirical studies on migration, social contexts – especially the characteristics of origins – often systematically affect individual behaviour and outcomes (see, for example, Windizio 2008). In this study on wages among migrants, although a set of village attributes can be controlled for, it is not possible to build an exhaustive model since many of the village-level factors cannot be directly measured or observed.

Fourth, if the wage and the network size are measured in the same year, the temporal order between the dependent and independent variable can be ambiguous. This is especially so when the higher wages of fellow villagers may attract more villagers to migrate, leading to a simultaneity bias when estimating the effects of networks on wages using cross-section data.

As long as any of the above four problems remained an issue[24], the standard OLS estimate could be biased and the estimated network effects on wages can be suspected to be spurious. If so, one cannot make any causal inference from the empirical results. In this study, I combine two identification strategies together to correct for the estimation bias. First, the Heckman's two-stage approach is employed to correct for the potential sample selection problem. Second, the information on natural disasters is exploited to develop a valid IV to deal with the more general endogeneity problem.

24 There might be other conventional sources of endogeneity, such as a measurement error, which normally attenuates the estimated effect down toward zero

3.2 Heckman's Two-stage Approach[25]

To address the bias resting on the unrepresentative sample, this chapter uses the Heckman two-stage approach (Heckman 1979, Heckman and Robb 1985). The procedure of Heckit estimation is as follows: in the first step, using all observations (in this study the whole sample including both migrants and non-migrants), a probit or logit selection model (in this study the model of migration decision) is estimated. Then, the inverse Mills ratio (or hazard ratio, the estimated probability of non-migration) is computed for each observation. The second step is to use the selected sample (in this study the migrants) and fit the substantive model where the inverse Mills ratio is also included as a covariate.

To practically distinguish the sample selection from a mis-specified model, however, the explanatory variables in the wage model should be a strict subset of those of the selection model. More importantly, at least one variable that does not appear in the wage model must be included in the selection model as an exclusion restriction (Wooldridge 2002, 2006, Cameron and Trivedi 2005). In this study, the size of the household labour force (i.e. the number of household members who are aged between 16 and 60) is used as the exclusion restriction.

25 Also known as the "Heckit" method.

That is, I assume that the size of household labour force influences the individual's decision to send its members to urban areas[26], while it does not affect the wage at the destination.

3.3 Natural Disaster as the IV

Although Heckman's two-stage method is useful for correcting for sample selection bias, it cannot deal with other types of endogeneity such as conventional omitted variable bias or the simultaneity problem. This is a problem, as the cross-sectional data to be used in this study does not provide any destination information. As a result, naively fitting the OLS model is very likely to lead to estimation bias. In the present study, what is needed is an IV that influences the wage at the destination only by affecting the size of the individual's village-based migrant network, and which is also uncorrelated with the error terms. This study exploits the richness of the data from the 2002 Chinese Household Income Project Survey (hereafter CHIPS 2002) and uses the intensity of any natural disaster in the origin village as an IV. Under certain assumptions, the natural disaster can be seen as an exogenous shock to the number of outflow migrants from the village, because the occurrence and destructive power of a natural disaster in a certain village is random.

26 For further details on how the household labour force positively affects migration, see Chen et al. (2008) and Zhao (2003).

The natural disaster is defined as the consequence of a natural hazard (e.g. drought, flood, earthquake), namely, the consequence of "extreme events that originate in the biosphere, lithosphere, hydrosphere or atmosphere" (Alexander 2000:9). There have been many scholarly explorations on the association between migration behaviour and the environment, and the underlying mechanism is intuitive (see Hunter 2005 for a review). McLeman and Smit (2006) argue that migration can be seen as an adaptation to climate change. Feng et al (2010) report find that a 10 per cent decrease in agricultural production in Mexico due to drought is associated with a 2 per cent rise in Mexican migration to US. In the context of the present study, the relationship between natural disaster and the migration behaviour could be generalized to a proposition which states that, the higher the destructive power of the natural disaster is, the more villagers are likely to move to urban areas during the subsequent months[27].

Natural hazards are common in China. Around 70 per cent of rural households are subject to natural hazards annually (see e.g. Wang and Zheng 2007) and the relationship between natural disasters and rural household income is very close (Gong and Zhang 2007). Importantly, in 2002, major natural disasters occurred in the first half of the year in China. For example, from January to March 2002, drought hit 24 provinces and deprived 15.9 million rural residents of drinking water. From April to May, strong rainstorms decimated the middle and lower

27 Given the huge income gap between urban and rural China, it is comfortable to assume that people are more likely to move to cities rather than to stay home after a disaster to repair the damage.

reaches of the Yangtze River[28]. Consequently, one has good reason to conjecture that the number of village migrants in 2002 should be correlated with the destructive power of the natural hazards that affected the village. As a matter of fact, this correlation can be statistically tested by the first-stage regression of the Two-Stage Least Square estimator.

It should be noted that losses caused by natural disasters can be region-specific, as the frequency and intensity of natural disasters varies across different regions and the capability of regions to overcome disasters also varies. This implies a possible relationship between natural disasters and wages, because wages at the destination can also be related to origins if migrants from a certain sending-area tend to move to a certain host area. However, if relevant regional dummies can be controlled for, natural disasters in a very small tract such as a village in a given year can be comfortably seen as random. In this study, I will control for regional dummies based on the Index of Regional Vulnerability to Natural Disaster by Wei et al (2004). In the robustness check, I will also add the dummies of origin provinces.

To be even more cautious about the exogeneity, the working assumptions are set up in such a way as to make the IV estimates as reasonable as possible: (i) Natural disaster in the origin village is assumed to not affect any labour market

28 The data are obtained from the 2002 Report on the State of Environment in China, by the Ministry of Environmental Protection of the PRC

condition at the destination. (ii) Omitted village-level factors are explicitly assumed to not be influenced by natural disaster within one year. (iii) Unobserved individual heterogeneity (e.g. earnings ability, reservation wage, labour market condition in destination cities) is assumed to be uncorrelated with the intensity of any natural disaster within one year. This is especially important, because a natural disaster may lower the reservation wage of a potential migrant. Given the huge income gaps between rural and urban China, however, even the lowest wage levels in cities are still attractive to villagers. This makes the reservation wage argument inapplicable in this setting.

3.4 OLS, Heckit and IV Models

Building on the preceding discussion about model choice and identification strategies, three different types of wage models are to be estimated, namely, the OLS, Heckit and IV-Heckit models.

First, the OLS model can be expressed by

$$W_{ig} = \beta_0 + \beta_1 S_g + \beta_2 X_{ig} + \beta_3 V_g + \epsilon_{ig} \qquad (1)$$

where W_{ig} is the wage income of the migrant, and i is the individual and g the village. S_g denotes the size of the migrant network, X_i denotes a vector of

individual factors, and V_g is a vector of observed village-level factors. ϵ_{ig} is the combined effects of unobservables at individual, village and destination levels. Thus, S_g is the major explanatory variable of interest and β_1 denotes the network effect. Notice that one of the most important premises for obtaining unbiased OLS estimates of β_1 is that Cov $[S_g, \epsilon_{ig}]$ =0. However, these two assumptions do not hold in the context of this study because of the potential presence of sample selection and other conventional endogeneity problem. As a result, alternative models are needed.

The Heckit model is:

$$W_{ig} = \beta_0 + \beta_1 S_g + \beta_2 X_{ig} + \beta_3 V_g + \beta_4 \widehat{P}_{ig} + \epsilon_{ig} \qquad (2)$$

where $\widehat{P}_{ig}$ is the inverse Mill ratio, a transformation of the predicted individual probability of migration calculated from the following selection equation (probit estimation):

$$P_{ig} = \gamma_0 + \gamma_1 F_{ig} + \gamma_2 S_g + \gamma_3 X_{ig} + \gamma_4 V_g + \mu_{ig} \qquad (3)$$

where F_{ig} is the household labour force which is not included in Eq. (2). As noted before, the explanatory variables in Eq. (2) should be a strict subset of those of Eq. (3).

Finally, based on the Heckit model, the IV model can be written as the following equation system：

$$W_{ig} = \beta_0 + \beta_1 \hat{S}_g + \beta_2 X_{ig} + \beta_3 V_g + \beta_4 \widehat{P}_{ig} + \epsilon_{ig} \qquad (4)$$

$$P_{ig} = \gamma_0 + \gamma_1 F_{ig} + \gamma_2 N_g + \gamma_3 X_{ig} + \gamma_4 V_g + \mu_{ig} \qquad (5)$$

$$S_g = \gamma_0 + \gamma_1 N_g + \gamma_2 X_{ig} + \gamma_3 V_g + \gamma_4 \widehat{P}_{ig} + \eta_g \qquad (6)$$

where N_g is the intensity of the natural disaster and η_g is the random error term in the equation denoting how the network size is influenced. Obviously Eq.(6) is the first-step regression of the IV estimation so that the inverse Mill ratio $\widehat{P}_{ig}$ as an exogenous variable is also included. In Eq. (5), given that the explanatory variable S_g is endogenous, it is not included. However, the instrumental variable N_g is included in Eq.(5) because “unless we have very reliable prior information, all exogenous variables should appear in the selection equation”(Wooldridge 2002:569)

4. DATA AND VARIABLES



Figure 2. Surveyed Provinces in CHIPS 2002

The CHIPS 2002 was jointly conducted by the Rural Survey Group of the NBSC and the Institute of Economics at the Chinese Academy of Social Science (CASS). The questionnaires were collected in February 2003 during the Lunar New Year. The advantage of doing so is that almost all outward migrants returned home to be reunited with their family members. The survey thus captures information on

all the members of rural households. The CHIPS 2002 gathered detailed information on demographic, household and village characteristics of the rural respondents. The data used in this study are from a sample of 37,969 individuals from 9,200 households, 961 administrative villages, 121 counties and 22 provinces[29]. The surveyed areas are grey in Figure 2[30].

4.1 Definition of Migration

Different criteria for defining migrant populations have been used in different national surveys in China (Liang 2007). In some studies, an individual is defined as a migrant if he or she crosses two administratively defined boundaries, while in others the definition is based on the duration of stay at the destination. The CHIPS 2002 asked the respondent to report the days away from their family in 2002 as well as the type of destination. The precise questions read: "How many days did you live outside this household in 2002?" and "Place of non-agricultural waged work: 1) within native village; 2) out of native village, within native township; 3) out of native township, within native county; 4) out of native county, within native province; 5) out of native province". Exploiting this information,

29 The sampling-frame for CHIPS 2002 is a sub-sample of the official rural household survey conducted by the NBSC. The administrative village, as opposed to the natural village, refers to a village which is legally recognized and numbered accordingly by the government. The result is that one administrative village may consist of more than one natural village.

30 Although most of this chapter is concerned with rural-to-urban migrants, it is helpful at the outset to show the difference between migrants and non-migrants. In Table 1 below, the first column describes the characteristics of migrants while the second column describes non-migrants.

this study sets the duration criterion as 180 days[31] and destination criterion as being away from the township where the migrant was born[32]. This is consistent with the criteria used by the NBSC and previous migration studies (e.g. Zhao 2003; Chen et al. 2008). Thus, migrants are defined as individuals (non-students) with "agricultural" *hukou* status who lived outside of their native-township for more than 180 days in 2002.

This definition may seem a bit arbitrary, but other definitions of migrant can and will be employed in robustness tests to examine whether the empirical results are sensitive to alternative definitions. In addition, all long-term out-of-township students, respondents younger than 16, and respondents older than 60 are dropped from the sample. Finally, due to some other variables of interest with missing values, the effective sample size ended up comprising 20,040 individuals from 8,308 households, 919 villages and 121 counties. Among them, there were 2,361 migrants from 1,787 households, 562 villages and 112 counties.[33].

4.2 Dependent Variable

The dependent variable is the daily wage of migrants in 2002 at the destination.

31 Since 1997, a six-month duration has become the standard criteria for defining migrants and was used in the recent Chinese Population Census by the NBSC (Liang 2006)

32 Surveys show that nearly 90 per cent of migrants live and work out-of-township (ROSC 2005).

33 Notice that the non-response rate on wages is around 10 per cent, which is common for working with income data.

This is calculated by the reported annual wage divided by the days working at the destination. There is only one migrant who had been totally unemployed throughout the days at the destination, so I dropped him from the sample. Due to the malpractice of local government and loopholes in the legal system, the problem of wage arrears among migrant workers has been severe in China, triggering social unrest and violence (see Li 2008; Hannan 2008). This explains why one can find in Table 1 that the minimum waged income is actually zero. In the sample, 274 out of 2,361 migrants went home empty-handed in 2002, suggesting that more than 8 per cent of migrants suffered from defaulted wages. As Table 1 shows, the average daily wage of the migrants is 14.38 RMB (*yuan*) and the maximum wage is 327 RMB. These are in general consistent with some previous studies (Li 2008; Hannan 2008)[34].

4.3 Independent Variable

The size of the migrant network is the explanatory variable of central interest[35]. As noted earlier, the measure of network size to be used is not based on administrative boundaries at the destination but on the origin villages. Migration studies (e.g. Munshi 2003; Rainer and Siedler 2008) often use this strategy, as most network relationships among immigrants at host areas are based on

34 The conservative estimation of the average monthly wage of migrant workers in 2002 is around 500 RMB and the optimistic estimation is around 700 RMB.

35 Here I have not used the proportion of migrants in a village as the explanatory variable because it is the number of fellow migrants that matters for individual outcomes at the destination. This is especially true since village populations range from less than 200 to nearly 9,000.

kinship, friendship, and especially belonging to a common origin place (e.g. Massey et al. 1987).

Table 1: Selected Descriptive Statistics

	Migrants *(N=2361)*				**Non-Migrants** *(N=17658)*			
	Mean	**Min**	**Max**	**S.D.**	**Mean**	**Min**	**Max**	**S.D.**
Individual Attributes								
Daily Wage (¥)	14.38	0	327	15.28	--	--	--	--
Working days	255	2	365	71.78				
Age	26.97	16	60	8.24	38.23	16	60	11.85
Age squared	795.15	256	3,600	530.58	1,601.91	256	3,600	902.86
Experience	3.99	0	38	3.57	2.00	0	43	4.75
Experience squared	28.72	0	1,444	71.52	26.61	0	1,849	102.70
Schooling years	8.14	0	16	2.19	7.02	0	16	2.67
Household labour force	4.81	2	11	1.38	4.44	1	12	1.35
Village Attributes								
Migrant Network Size	252	0	1,530	251	153	0	1,530	189
Average Income Per Capita 1998	1,858.5	0	13,980	1,027.9	2,079.78	0	13,980	1183.2
Labour force 2002	1074	65	4813	703	954	42	4813	647
Distance from county seat (km)	27.14	1	140	20.66	24.27	.5	160	21.l1
Intensity of Natural Disaster	.11	0	.85	.13	0.12	0	0.9	.16
	Freq.		**per cent**		**Freq.**		**per cent**	
Individual Attributes								
Male	1,479		62.64		9,233		52.29	
Female	882		37.36		8,425		47.71	
Village Attributes								
Suburb of large/middle city	60		2.54		1,372		7.77	
Non-suburb	2,301		97.46		16,286		92.23	

Using village-level migrant networks rather than networks at the destination based on common origin is particularly useful in China. As mentioned in Section 2, labour migrants in China return home frequently during a year. According to the data to be analysed, only 28 per cent of migrant villagers reported that they had not talk much about job-related information when they are in villages. These make the village of origin a pivotal site for facilitating the spread of job leads among fellow villagers even at different locations. Furthermore, using origin community as the social unit to construct the measure of migrant network does have its advantages: it offers the opportunity to instrument the explanatory variable using any natural disaster that hit the village.

Still, one may argue that, since the daily wage and the network size are measured in the same year, the mutual relationship between the dependent and independent variable leads to a simultaneity bias when estimating the effects of networks on wages. However, this is hardly an issue in the setting of this study for two reasons. Firstly, most migration decisions are made in the first half of the year, especially right after the Lunar New Year in early February. As a result, the total outflow migration of the village in February or March is basically the same as the total outflow of the whole year. Secondly, even though there might be mutual causality between the wage variable and the network variable, using natural disaster as the IV can deal with this problem because the simultaneity

bias is in essence a special type of omitted variable problem[36].

The measure of network size is obtained from the responses of the village heads to a question posed in the survey. Specifically, all the village heads were asked about the number of outflow migrants in 2002 from his or her village. The precise question is "Approximate total number of out-migrants working out of township more than six months". As is shown in Table 1, the average number of outflow migrants is 252, and the largest migrant network contains 1,530 migrants while the smallest is zero, which means the respondent is the only migrant in his or her village.

4.4 IV and Regional Dummies

The destructive power of the natural disaster that serves as the IV is measured by the percentage reduction in agricultural output in the village. For each village surveyed by CHIPS 2002, the village head was asked to answer the question "Was there any natural disaster during the year 2002" and the follow-up "If yes, what was the percentage reduction in agricultural output as compared to an ordinary year?" For those villages without the presence of any natural hazard in 2002, the variable is coded as zero, showing that no loss was induced by the natural

36 Obviously it looks more reasonable to use the network size of the last year (2001) to predict the individual wage in 2002. However, the data does not provide us with the information regarding natural disasters in 2001, so that I cannot correct for the estimation bias using the network size of 2001 as the independent variable. However, I did find a positive correlation between the network size of 2001 and the wage of migrants in 2002 using the standard OLS model.

disaster. The village heads who could not provide accurate figures were asked to provide rough estimates for the percentage reduction in agricultural output. As can be seen in Table 1, the average percentage reduction in agricultural output among migrants is around 11 per cent, with a maximum percentage of 85 per cent.

As discussed before, loss by natural disaster can be region-specific and therefore relevant regional dummies need to be controlled for to make sure the IV is exogenous. For a given region, it is easy to see that the agricultural loss by natural disasters is determined by the overall capability of the region to overcome disasters rather than merely the frequency and intensity of natural disasters within the region. Considering this, I control for the dummy of the Index of Regional Vulnerability to Natural Disaster which is created by Wei et al. (2004). Using a data envelopment analysis-based mode, Wei and his colleagues generated a five-level category to rank the regional vulnerability to natural disaster. This is calculated based on the total cost of the damage, the density of population, the GDP, and the number of people affected by disasters from 1989 to 2000. The category is illustrated in Figure 3. Notice that from level 5 to level 1, the capability to overcome disasters increases. Finally, given that around 28 per cent of the migrants migrated within provinces in this study, the wage differences across 22 sending-provinces may be considerable. I will include the provincial dummies in the robustness check.



Figure 3. Five Levels of Regional Vulnerability to Natural Disaster

4.5 Other Controls

The individual-level controls include the migrants' age, age squared, gender, schooling years, experience, experience squared, type of occupation, and region dummies. As can be seen in Table 1, the mean age of the migrants is around 27. Most of the migrants are male and females only account for less than 37.5 per

cent. Married migrant workers account for 42.6 per cent of the whole migrant population in this data. When it comes to educational attainment, the mean years of schooling among migrants is eight years. Note that since it is the off-farm experience that matters for the waged income at the destination, the experience variable is not calculated in the standard way, such as in the Mincer Equation which is widely used in econometrics literature (i.e. age minus years of schooling minus six). Rather, it is measured in this study by how many years migrants have worked in cities.

Although not reported in Table 1, rural-to-urban migrants are engaged in 18 types of occupations including "agriculture (planting, herborization)"[37], "forestry (husbandry, fishing or water conservancy)", "mineral and geological survey and prospecting", "industry", "construction", "transportation, communications, posts and telecommunications", "commerce and trade", "restaurants and catering", "materials supply and marketing, warehousing", "real estate", "public service", "personal services or consulting services", "public health, sports and social welfare", "education, culture, arts and broadcasting", "scientific research and professional services", "finance and insurance", "government agents, party organisations and social groups" and "others". Among them, more than 60 per cent of migrants are engaged in the "industry" and "construction" sectors.

37 There are a small number of peasants moved to another place (e.g., suburb of a provincial capital) and were engaged in agriculture-related industry.

The village-level controls include the per capita net income (including both migrants and non-migrants) in 1998, the distance to the closest county seat, a dummy denoting whether the village is a suburb of a large or medium-sized city or not, and the number of labour force of the village. Including the lagged mean income of the village is an attempt to proxy for unobserved village-level earning ability. Notice that this method of controlling for unobserved factors is far from satisfactory, since it may introduce more noise into the model. The distance to the closest county seat and being a suburb of a large or medium-sized city both proxy for access to beneficial job leads and certain unobserved individual abilities which are suspected to influence the wage in a direct way. The variable of village labour force is used to control for the potential network size of migrant and the population of the village.

5. ESTIMATION RESULTS

The estimates obtained from the OLS, Heckit and IV-Heckit models are all presented in Table 2 to provide comparison values. Given that respondents are actually "clustered" by villages, the observations of this study may not be independent. I thus use the robust cluster variance estimators for the three models so that the standard errors here are all robust to within-cluster correlation. In addition, it is worth highlighting that, when fitting the model, the wage appears in logarithm form given that the skewness statistic for the waged

income is 4.46, indicating a considerable right-handed skewness. For the same reason, the logarithm of the network size is used.

The first column of estimation results in Table 2 contains the OLS estimates. The coefficient of the network size variable is around 0.125 with a robust standard error of 0.0349, suggesting that migrants who came from a village with around an 8 per cent larger migrant network (outflow migrants) are associated with 1 per cent higher wages. That is, a 1 per cent increase in the network size would produce an expected .125 per cent increase in wages, net of other factors. While the age and age squared variables both significantly influence wages at the 0.01 significance level, the directions of the two effects are different. This implies a diminishing marginal effect of age. Experience of working in cities is also an important determinant of wages, yet its squared value is significant at only the 0.1 level, providing only weak evidence of a diminishing return to experience.

Also as expected, years of schooling do not seem to be related to the wages of the migrant workers. This could be due to fact that the occupations that migrants take in contemporary China are extensive but mainly low-skill and the average educational requirement is quite low. What is unexpected is the role of gender. According to the empirical result, the wage difference between men and women is not significant. This is probably because of the limited occupation choices, as well as the ceiling effects in terms of promotion among rural-to-urban migrants.

On the one hand, they are mainly concentrated in low-income jobs in which the gender difference in wage is marginal all else being equal; on the other hand, the opportunities for them to be promoted within workplaces is very limited, so that male migrants do not have more advantages than women. Finally, in general the village characteristics do not affect wages significantly.

The second column presents the Heckit estimates. When selection in migration using unobserved ability is explicitly taken into account, the network effect increases from 0.125 to 0.263, and remains highly significant. The partial effect of the inverse Mills ratio, i.e., the predicted probability of migration, is around 0.75 and significant at the 5 per cent level. As for the other controls, compared with the estimates from the OLS model, the effects of age seemly disappeared[38], while the role of experience increases and the influences of the size of the village labour force and the village location become significant. These findings strongly suggest that there is considerable selection based on unobserved individual ability when people make migration decisions. Furthermore, if we believe that more able people benefited less from the origin-based networks, the larger Heckit estimate of network effects reveals that the OLS estimate may have largely reflected the network effects among more able people who tend to migrate.

38 More precisely, the effects of age did not disappear. Rather, because the coefficient for age is not significant, it means that we cannot reject the hypothesis that the parabola described by the squared term is symmetric relative to the x-axis. The significant squared term implies that the effect of age on income is curvilinear; whether it goes up at a diminishing rate or down at an increasing rate depends on where your data points lie relative to the maximum value of the parabola.

Table 2: OLS, Heckit and IV Models of Migrants' Wages (Logged wages as the dependent variable N=2361)

Independent Variables	OLS	Heckit	IV
Logged network size	.125***	.263***	.628***
	(.0349)	(.076)	(.232)
Age	.06***	.0352	.0456
	(.0197)	(.0238)	(.0282)
Age squared	-.0009***	-.001***	-.0009***
	(.0003)	(.0003)	(.0003)
Gender	.0204	.0458	.0503
	(.0449)	(.0466)	(.0535)
Years of schooling	.008	.0182	.0204
	(.0126)	(.0137)	(.0159)
Urban experience	.0588***	.0993***	.0904***
	(.0181)	(.0276)	(.0170)
Urban experience squared	-.00145*	-.00351***	-.00357***
	(.0008)	(.0013)	(.0017)
Per cap. village income 1998	.0194	-.0471	-.0845
	(.0397)	(.0434)	(0.165)
Village labour force	-.00007	-.0001**	-.0004**
	(.00006)	(.00006)	(.00019)
Distance to closest city	.0014	.0047**	.0013
	(.0015)	(.0022)	(.0030)
Suburb of large city	-.992	.469*	-.138
	(.212)	(.279)	(0.454)
Occupation dummies	YES	YES	YES
Region dummies	YES	YES	YES
Constant	.192	-1.566	-1.746
	(.651)	(1.150)	(1.404)
Inverse Mills ratio		.754**	.353**
		(.379)	(0.147)
First stage F-statistic on IV			*17.42*
Hausman test of endogeneity			*p= 0.007*

Note:

i) Standard errors are in parentheses and robust to heteroscedasticy and within-village correlation.

ii) * p < =0.1 ** p <= 0.05, *** =p < 0.01

Of interest is why more able people benefit less from village-based migrant networks than less able migrants? This may be explained by the fact that migrants with a higher earning ability are also more able to initiate social ties with urban communities at the destination. Thus, their job-searching relies less on village-based networks since the new ties built at the destination generate more valuable job leads. For example, Qi (2007) and Xu (2009) both report that some migrants start to build up social ties with urban communities at the destination. Cao (2003) reports that migrants with urban-based social ties (established after migration) send back more remittances and have more savings than those without new friends in cities. He also finds that the size of newly established urban-based networks is positively associated with labour market outcomes. These findings show that urban ties are more beneficial than origin-based ties in matching migrants to jobs, implying that less able migrants with fewer new ties may rely more on the origin-based network than high-ability migrants.

Several important results are found in the third column which contains the IV-Heckit estimates. First of all, the quality of the IV is good since it comfortably passed underidentification and weak-identification test. As is shown at the bottom of the third column, the F-statistic for significance of the IV in first-stage regression of the endogenous network size is 17.42 and significant at 0.05 alpha

level[39]. In light of Stock and Watson's (1997) rule of thumb that the first-stage F-statistic should be at least 10 for a single endogenous regressor, this implies a significant statistical association between the natural disaster and village migration. It is noteworthy that the F-statistic is also larger than Stock and Yogo's (2005) weak IV critical value (16.38 for 10 per cent maximal IV size, one endogenous regressor and one IV), suggesting that the relevance of IV is sufficiently strong to reject the weak IV hypothesis with less than 10 per cent relative bias toleration. This makes us fairly confident in the IV estimates. More importantly, the Hausman endogeneity test indicates that we can reject the null hypothesis that there is no systematic difference between the estimates from the OLS and the IV model. Since the IV estimator always yields consistent estimates given a valid IV, coefficients estimated from it should be more meaningful. This further suggests that the size of migrant network is indeed endogenous.

Compared to the Heckit estimate, the effect of network size obtained from the IV model is substantially higher. It increases from less than 0.263 to around 0.628, remaining highly significant given the 1 per cent significance level. As for other controls and the inverse Mills ratio, no major changes are found in the partial effect except that the role of the location of villages disappeared. Consider two villages sending migrants to the same city. All else being equal, a 10 per cent increase in the size of the network implies an expected 6.28 percent increase in

39 The results of the first-stage regression can be seen in the Appendix 1

wages. However, an issue arises that the larger IV estimate of the network effects looks counter-intuitive. This is because one would expect that correcting for the rest of the estimation bias would lead to a mitigated IV estimate in comparison to the Heckit estimate. This will be discussed in the following section.

Before proceeding to the discussion of the IV estimate, the results should be examined to see whether they are qualitatively robust with respect to changes in model specification, the definition of migration and differently truncated samples (due to wage outliers). Firstly, I used the alternative criteria of destination location and duration for migrants to test whether the same pattern of network effects on wages can be found. To this end, migrants living out-of-native-village for more than 180 days, out-of-native-county for more than 180 days, out-of-native-township for more than 90 days and out-of-native-township for more than 240 days are examined respectively. Secondly, I fit models including province dummies of origin. Thirdly, considering the potential outliers in terms of wages, I tried deleting observations in the top 99 per cent, the top 95 per cent, the bottom 99 per cent and the bottom 95 per cent quantiles from of the sample. The sample without those who report a zero-wage is also used to see whether the analysis is sensitive to zero-wages. As can be seen in Table 3, all of these estimates show similar patterns in comparison with the estimates in Table 2, suggesting that the estimates in Table 2 are robust. Although not reported here, I also explored interactions between gender and other personal attributes and the

estimates are similar and stable.

Table 3 Estimates from Robustness Check

	Definition1	***Definition2***	***Definition3***	***Definition4***
Destination	Out of village	Out of county	Out of township	Out of township
Duration	>180 days	>180 days	>90 days	>240 days
OLS estimate	**.151*** **	**.124*** **	**.096*** **	**.139*** **
Heckit estimate	**.214*** **	**.270*** **	**.154*** **	**.261*** **
IV estimate	**.773*** **	**.619*** **	**1.058*** **	**.829*** **
N	2486	2135	2800	1975

	With Provincial dummy	***Top 99 per cent deleted***	***Top 95 per cent deleted***	***Zero-wage deleted***
OLS estimate	**.073***	**.137*** **	**.143*** **	**.0722*** **
Heckit estimate	**.121** **	**.194*** **	**.186*** **	**.252*** **
IV estimate	**.403** **	**.605*** **	**.731*** **	**.468*** **
N	2361	2359	2254	2102

Note: i) * p < =0.1 ** p <= 0.05, *** =p < 0.01

6. INTERPRETATION OF THE IV ESTIMATE

The IV estimate suggests that network effects obtained from conventional Heckit model are downwardly biased. As mentioned before, this is intriguing because correcting for the partial effects of individual heterogeneity and unobserved

village factors should in theory result in a smaller estimate. One possible reason for this is that the omitted variable is positively correlated with network size yet negatively correlated with wages, holding the inverse Mills ratio and other observed factors constant. Notice that to this point the effect of migrant networks has been assumed to be constant across individuals and villages. However, this assumption may be too strong in the context of rural-to-urban migration. It thus seems reasonable to interpret the IV estimate within the framework of heterogeneous network effects.

If natural disasters influence migration mainly among villages/individuals benefiting from more extensive networks, the IV estimate would be higher than the Heckit estimate. This is simply because the Heckit estimate applies to all villages/individuals while the IV estimate mainly applies to the sub-group of villages/individuals more affected by natural disasters. With the presence of casual effects heterogeneity, the IV estimate can be interpreted as a Local Average Treatment Effects (LATE) (Imbens and Angrist 1994; Angrist et al 1996)[40]. For example, Card (1995) uses college proximity as an IV to identify returns to schooling. He reports a larger IV estimate than a conventional OLS estimate, and accounts for it by arguing that college proximity affects schooling mainly among those children who have poorly educated parents and thereby benefit more from education.

40 Importantly, interpreting IV estimates under the LATE framework should be done on the condition that the relationship between the IV and the explanatory variable is one of monotonicity, i.e. the natural disaster should influence village migration in one direction.

If the IV estimate is to be interpreted as a class of LATE, we must ask about the mechanism that explains how natural disasters influence migration and why network effects differ across villages/individuals. One possible mechanism is that that less able people (in terms of earning ability at the destination) are more responsive to natural disasters, since they have a relatively lower ability to compensate for losses due to natural disasters. That is, villagers of lower earning ability are more likely to be "pushed" out from the villages by natural hazards. If this is the case, the IV estimate can be interpreted as a weighted average network effect and the weight for less able migrants is relatively higher. Since it is clear that less able migrants benefit more from origin-based networks, the IV estimate mainly captures the network effects among them and it would then be higher than the Heckit estimate.

Another interpretation can be that villages with less migration tradition are more responsive to natural disasters and migration tradition is associated with certain unobserved individual ability. For example, a higher level of migration tradition entails a strong norm or pressure of migration. Given this class of deeply rooted social norms, regardless of being hit by natural disasters or not, the momentum of outward migration should be relatively stable in villages with a stronger migration tradition. As a result, exogenous shocks from natural hazards could lead to considerable changes in the size of the migration population for villages

with weaker historical migration. In addition, members in villages with weaker migration tradition may be relatively less able as compared to their counterparts in villages with stronger migration tradition. Especially, one could expect that a weaker migration tradition should be associated with less access to new information and technology from urban areas, and less human capital from return migrants, holding other things constant.[41]

Aside from being correlated with lower earning ability of villagers, weaker migration tradition can also be associated with larger network effects. For example, weaker migration tradition can be suspected to render smaller migrant networks since migration in the origin communities can be path-dependent. Suppose the marginal effects of network size on wages are diminishing, migrants from villages with weaker migration tradition then would benefit more from the migrant network. This is very likely the case in China because almost all of the migrants are channelled into low-skill occupations characterized by abundant job vacancies and low wages. Thus, below a certain threshold size of networks, the network effect for finding a "good" job can be considerable because a bigger network entails more job leads. Yet when the network size exceeds this threshold, the return to network size may be declining since in his/her wage ranges, it is the quality rather than the quantity of job leads that matters.

41 Still, one may argue that villages with better living conditions or higher productivity may have lower migration tradition due to a lower level of motivation. That being said, given the huge income gap between rural and urban China, it is not unreasonable to assume that the "desire" of migration does not vary between villages.

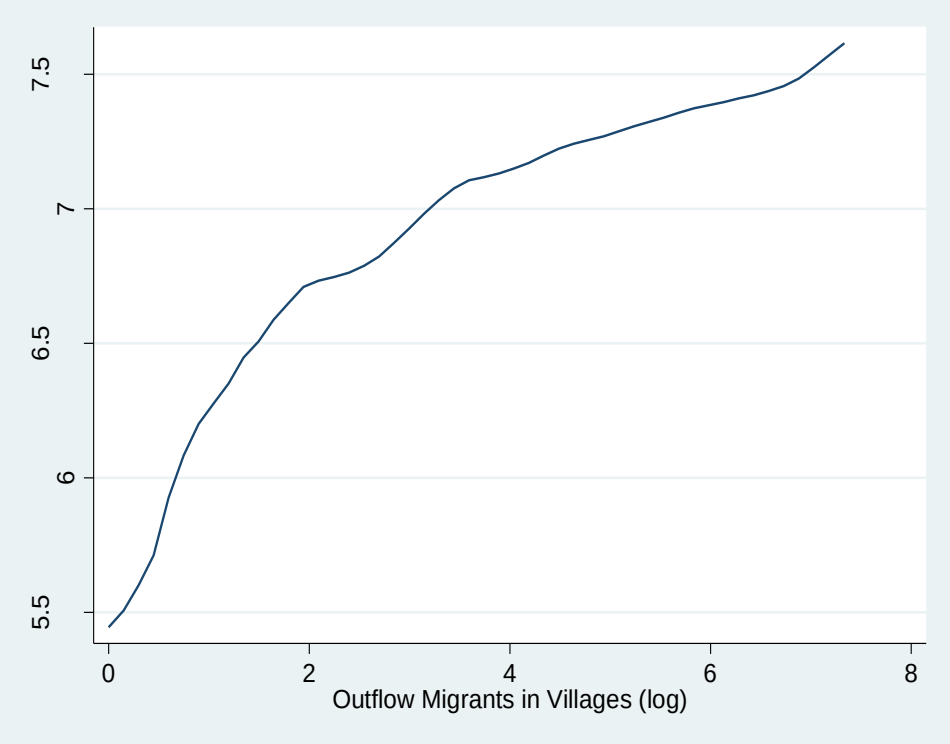


Figure 4. Local Polynomial Smooth Curve
showing diminishing marginal effects of network size on wages

The models reported in Table 2 and 3 assumed that the slope of the network size variable, the marginal effect of networks, is fixed. If this assumption is relaxed, decreasing marginal effects can explain why migrants from villages with a weaker migration tradition experience larger network effects. To illustrate this, a local polynomial smooth curve is used to estimate the effects of network size on wages[42]. As is shown in Figure 4, with the logged network size going beyond 3.0, the slope of the curve substantially decreases. This implies that when the network size goes beyond 20 (i.e., exponent of 3), the network effect substantially declines. This pattern is in general consistent with the speculation

42 The local polynomial smoothing fits simple models to localized subsets of the data to build up a function that describes the deterministic part of the variation in the data. The advantage of local polynomial smoothing is that it does not require the specification of a function to fit a model to all of the data in the sample. This is useful when we do not know exactly the relationship between two variables.

that the marginal effect of migrant networks is not fixed but rather diminishing.

7. CONCLUSION

Rural-to-urban migration in China since the late 1970s represents the largest labour flow ever observed in the world. Despite the proliferation of research seeking to identify the mechanisms and measure the magnitude of internal migration, little emphasis has been placed on probing the direct causal effects of migrant networks on labour market outcomes at the destination. This chapter explores the causal effects of the size of migrant networks on wages among migrants in China. It complements recent research on the labour market effects of migrant networks in market economies such as the US and Germany (e.g. Munshi 2003; Rainer and Siedler 2008). The major innovation of this chapter is the focus on the role of the total outflow of migrants at the origins, and the combination of Heckman's two-stage method and IV approach.

Controlling for the unobserved factors influencing the migration decision, identification is achieved through instrumenting the network size by the intensity of natural disasters occurring in the sending-villages of the migrants. The empirical results show that the size of the migrant network significantly improves the wages of migrants. The mechanism is straightforward:

village-based networks matter because these networks transmit job-related information in cities among migrants and hence influence wages at the destination. This is consistent with findings reported elsewhere and implies that the mechanism of job-search among migrants in the context of state socialism is similar to that in other contexts. Obviously, focusing on the role of the total outflow migrants at origin level does not deny the role of networks based on common origin at the destination.

One potential criticism of this study is that the validity of the IV approach cannot be statistically tested for. This might be one of the reasons why sociologists rarely use the IV approach. However, as Morgan (2002) puts it when he calls for more attention to be paid to the IV approach, "within the new literature on counterfactual causality, instrumental variables estimators have received substantial attention from both statisticians and econometricians … even fallible IVs can be useful for examining patterns of causal effect heterogeneity." Naturally, the results reported in this chapter should not be regarded as definitive. However, the results and the identification strategy improve our understanding of the mechanism though which cohesive migrant networks influence the wages of migrants.

Appendix 1

Estimates of the First-Stage Regression

Logged Network Size as the Dependent Variable (N=2361)

Independents Variables	
Loss by disasters (IV)	1.494***
	(.360)
Age	-.0263
	(.0253)
Age Squared	-.0001
	(.0003)
Gender	.0260
	(.0473)
Years of Schooling	-.004
	(.0155)
Experience	.0987***
	(.0334)
Experience Squared	.0987**
	(.0334)
Per Cap. Village Income 1998	-.0414
	(.125)
Village labour force	.00086***
	(.000076)
Distance to Closest City	.00838**
	(.00332)
Suburb of Big City	1.359**
	(.412)
Occupation Dummies	YES
Region Dummies	YES
Inverse Mills ratio	.819
	(.469)
Constant	1.173
	(1.302)

Note: i) * p < =0.1 ** p <= 0.05, *** =p < 0.01

Chapter 4

Do Friends' Diplomas Matter? Estimating Network Effects in the Presence of Selection

1. INTRODUCTION

This chapter focuses on the effects that network resources have on individuals' socioeconomic outcomes. In different research contexts, this class of network resources has been expressed as accessed social capital (Lin 1999) and network social capital (Mouw 2006).[43] Irrespective of the terminology adopted, the core idea is that the characteristics of networks, friends or groups can affect an individual's social-economic outcomes. These characteristics can be seen as useful resources embedded in social networks. They include information, influence or social credentials, and can be represented by wealth, power, rank, prestige or status in most societies (Lin 1999).

Social resource theory argues that resources embedded in social networks are beneficial for individual outcomes, as individuals can access and use the

43 Van der Gagg and Snijders (2003:200) provide a useful definition of this class of social capital as "the collection of resources owned by the members of an individual's personal social network, which may become available to the individual as a result of the history of these relationships".

resources possessed by others to gain advantages (Lin 1982, 1990).[44] A large number of empirical studies support this perspective (see Lin 1999 for a review). However, that said, the estimated effects of network resources in many previous studies may be biased, since social networks are often formed in a non-random way. For example, it has long been established that social interactions are often driven by interpersonal similarities, namely, social homophily (e.g. McPherson, Smith-Lovin and Cook 2001). According to Mouw (2003, 2006), a fundamental challenge in estimating the causal effect of network social capital lies in the difficulty of separating the "genuine" role of network resources from the effects arising from unobserved individual characteristics such as selectivity.

Using data from the 2003 Chinese General Social Survey (CGSS), this analysis assesses how friends' educational attainment affects one's labour force participation and income. Network resource here is measured as the average years of schooling of discussants of important matter (i.e. named friends within one's core discussion network). Given that the core discussion network is most likely to be selective, the paper uses an innovative Instrumental Variable (IV) method formalised by Bayer and Ross (2009). The core idea behind this method is to instrument for each individual's observed network resource with the average network resource of a set of observationally identical individuals. The

44 In the context of the labour market, although it is often not clear whether one has actually used network resources, those who have more informative or influential friends receive at least routinely better job-related information. The network resource, therefore, at least exerts its role as the invisible hands of social capital (Lin 2004, 2008).

empirical results demonstrate that the IV estimates of the network effects on employment and income are larger than those obtained from single-equation models, namely, the standard probit model and the ordinary least square (OLS) model. Because the construction of the IV is based on certain assumptions, the aim of this chapter is to point out and correct for selection bias rather than to precisely give the magnitude of the effects of educational network resource.

The significant and larger IV estimates imply that individuals with poor unobservables in terms of employment or income (e.g. less earning-ability, disadvantaged positions in labour markets etc.) have adopted a compensation strategy and tend to select friends with higher educational attainment. This compensation strategy suggests the presence of social heterophily in the formation of social networks. The finding here is consistent with Lin's (2008) position that, when the resources embedded in homophily-based networks are not sufficient for a specific goal, individuals may reach beyond the inner layer of their social networks to access heterophilous resources. This finding gives further weight to Mouw's (2003, 2006) call for taking seriously the selection problem when trying to identify the causal effects of network social capital.

The rest of the paper is organized as follows. Section 2 presents a brief literature review of social resource theory and the endogeneity problem plaguing causal inferences in estimating network effects. It also provides some background

information on the job-seeking processes in China. Section 3 describes the models, focusing on how to deal with the self-selection problem as a source of endogeneity bias. Section 4 discusses the data and variables. Section 5 presents empirical results from single-equation models and Section 6 presents the results from the IV estimators. Section 7 interprets the results and, finally, Section 8 summarises and concludes.

2. BACKGROUND

Although the theoretical intuition behind the idea that individuals access and use social resources is very appealing, to identify the causal effect of network resources is difficult when using observational data. The major concern here is that individuals may choose friends or groups according to a certain unobserved individual characteristic.[45] As a result, estimated effects from standard single-equation models (i.e. the OLS or the probit models) are very likely to be biased. Most previous studies in sociology, however, have not paid much attention to this problem.

2.1 The Selection Problem in the Literature

45 According to Mouw (2006), the other source of endogeneity bias in estimating the role of network social capital aside from self-selection is the simultaneity problem. This is not addressed in this study. However, as will be discussed in the section on data and variables, since network resources are measured by average educational attainment among one's friends, the simultaneity problem can be sidestepped.

The network effects on individuals' outcomes reflect various mechanisms. The most widely-discussed mechanism is that job leads can be transmitted within social networks and job-related influences can also be accessed and mobilised through social ties. There is an established literature supporting this view, comprehensively reviewed by Granovetter (1974:1995) and Lin (1999). For example, Lin and Dumin (1986) and Völker and Flap (1999) find that the "resource reachability" of a network (i.e. the highest occupational prestige score or educational attainment of contacts in the social network) is positively correlated with individual status in the US and the former East Germany. Sprengers,Tazelaar and Flap (1988) report that the weighted sum of the status of friends positively predicts individual status in Holland. In addition, there is a wide range of studies (e.g. Burt 1992; Meyerson 1994; Erickson 1996; Flap and Boxman 2001; Seibert et al. 2001) focusing on how the structure of social networks influences job outcomes.

Educational attainment can be seen as a typical indicator of one's human capital. It is therefore natural to suspect that the pool of close friends' human capital can be beneficial to one's social-economic outcomes. For example, Mouw (2003) finds a positive association between the respondent's wage level and the average educational attainment of his/her friends. However, as he points out, due to the potential selection problem, the findings related to the effect of network social

capital cannot be seen as causal. Selectivity is important for the study of network resources because it suggests a competing explanation to that reached in the findings of previous research. For example, this is especially true if people selectively socialise with those who are similar to themselves in terms of a certain unobserved ability and this ability also affects individual outcomes, meaning that the estimated network effects may just reflect the role of the shared ability rather than the role of network resources. If the selection problem cannot be addressed, the estimated network effects can be spurious.

Self-selection has long been a concern in social interaction studies, such as research on neighbourhood and peer effects, as reviewed in Manski (1993), Moffit (1995, 2001), Durlauf (2002), Mouw (2006) and Soetevent (2006). A set of identification strategies has been used to address the self-selection problem. For instance, the fixed-effects model has been widely used to rule out the effect of time-invariant unobservables from which selection can possibly arise (see e.g. Weinberg, Reagan and Yankow 2004, Hedström, Åberg and Kolm 2003). Other useful methods dealing with potential selection bias include modelling the selection process and natural experiments[46] (e.g. Korpi 2001; Oreopoulos 2002; Dujardin and Gpffette-Nagot 2007; Hellerstein et al. 2009). The IV approach has also been employed to correct for selection bias. Bayer and Ross (2009), for instance, instrument for each individual's observed neighbourhood

46 The propensity score matching method is not useful here, as it can only be implemented under the assumption that self-selection is totally caused by the observed variables.

characteristics with the average neighbourhood characteristics for a set of observationally identical individuals. Although these studies do not focus on friendship networks, they provide useful templates for dealing with the potential self-selection bias arising from individual unobservables in network effects studies.

2.2 Job Search and Network Resources in China

The important role of network resources in Chinese society has been well documented. For example, it has been widely accepted that *guanxi*, as a class of social ties characterised by reciprocal obligation, constitutes one of the core mechanisms of interpersonal interactions in China (Walder 1986; Hwang 1987; Bian and Ang 1997). When it comes to the job-matching process, Bian (1997) finds that Chinese people rely more on strong ties because strong ties are more efficient in obtaining help from job-assigning authorities given the dominant system of state job assignment. With regard to network effects on individuals' socioeconomic outcomes, Bian and Ang (1997) find that a job helper's status is positively associated with the attained status of the job changers. Lin (2003) also reports on the significant effect that contacts' supervising unit rank has on job-seekers' supervising unit rank. These pieces of evidence suggest that resources embedded in social ties matter for individuals' labour market outcomes in China.

More importantly, based on the 1988 Tianjin Study, Bian (1997) reports that 45 per cent of respondents used contact-chains (i.e. job-search chains involving a second intermediary/indirect linkage) to find a job. Moreover, the study also states that the ultimate helpers for contact-chains users are significantly more resourceful in comparison to the helpers for direct-tie users.[47] He also finds that the ultimate helper was usually targeted first and the search for the "right" intermediary was planned accordingly by the job-seeker. These facts imply that, under certain circumstances, people initiate social ties with, or, select those who are dissimilar to themselves in terms of social resources.

Also in a study focusing on the inequality of social capital, Lin (2003) reveals that women in China rely more on contacts to find jobs and have longer and stronger contact-chains than men. Lin (2003: 148) thus argues that "females and other racial and ethnic groups are disadvantaged and need to take extraordinary actions through cross-gender and cross-racial/ethnic ties to gain access to better resources in their instrumental actions". This suggests that disadvantaged people may adopt compensation strategies (e.g. initiating social ties with those with better resources) to gain advantages in the labour market. Of course, this phenomenon is not confined to China. According to the findings in Renzulli and Aldrich (2005), female business owners have greater occupational heterogeneity

47 The resource is measured by three variables: whether or not the helper holds an administrative position, the level of helper's work unit in the hierarchy of work units, and occupational status.

scores and activate more resources from others than male business owners in the US.[48]

Findings in Bian (1997) and Lin (2003) imply that the motivations behind and the process of relationship formation can be very complex. On the one hand, people tend to make friends with similar people in light of the principle of social homophily; on the other hand, they are also likely to socialise with those who are dissimilar to them in terms of resources. This is especially true when, as Lin (2008) remarks, resources embedded in homophily-based networks are not sufficient for a specific goal (to find a good job, for example), when individuals may reach beyond the inner layer of their social networks to get access to heterophilous social capital. This is called heterophilous selection in this study.

It should be noted that both Bian (1997) and Lin (2003) focus on the social resource embedded in a dyad (i.e. the resource of a given job helper) formed for a specific instrumental goal, which is getting a job. Thus, their findings merely suggest the potential presence of heterophilous selection in the forming of job-seeking ties which are obviously goal-specific and often temporary. Consequently, it looks unreasonable to conjecture that heterophilous selection may also arise in the forming of other types of social ties, such as core discussion networks. In particular, it is intuitive to expect that it is the principle of social

48. Renzulli and Aldrich (2005) report that there is no significant difference in activating resources between men and women, given the significance level of 0.05. However, according to their empirical finding, the difference is significant at the level of 0.1.

homophily that affects the formation of core discussion networks, because people with similar interests and backgrounds are more likely to develop intimate and emotional relationships.

However, as Bian (2008:81) argues, the formation of social networks among Chinese is a more conscious process in which "individuals develop social capital by cultivating, maintaining, and utilizing certain social relationships". More importantly, as I will show in the data description section, core discussion networks in China are very important sources of embedded resources for instrumental actions such as getting a job or achieving a better salary. As a result, one can even expect the presence of heterophilous selection in the formation of core discussion networks among Chinese. If so, controlling for individual unobservables which may lead to selection, the estimates of the effects of core discussion networks on individual outcomes would be higher than those obtained from standard single-equation models.

3. MODELS AND METHODS

In order to test the network effects on individual outcomes, I fit two types of model using the CGSS data: one for identifying the network effect on labour market participation and the other for examining how networks influence annual income.

Firstly, a probit model predicting labour market participation can be written as:

$$P(y_i = 1) = \Phi(\beta_0 + \beta_1 S_i + \beta_2 X_i) \quad (1)$$

where y_i =1 denotes that individual *i* is employed, S_i is the variable of network resource, the average years of schooling taken over the social network of the foci individual *i*, and X_i is a vector of observed individual characteristics as controls.

It should be noted that the underlying indicator function for Eq. (1) is:

$$y_i^* = \beta_0 + \beta_1 S_i + \beta_2 X_i + \varepsilon_i \ , \ y_i = 1 \text{ if } y_i^* > 0 \quad (2)$$

where ε_i is the error term and y_i^* is the latent variable. That is, y_i is one if $y_i^* > 0$.

Second, in much the same way, an OLS model predicting income can be expressed by:

$$I_i = \gamma_0 + \gamma_1 S_i + \gamma_2 X_i + \varepsilon_i \quad (3)$$

where I_i is the income of individual *i*.

However, the estimates obtained from Eq. (2) or Eq. (3) are likely to be biased, due to the fact that networks are likely to be selective and therefore S_i is likely to be correlated with the unobserved individual attributes ε_i.

Obviously what is needed is an IV that is uncorrelated with the individual outcomes but strongly correlated with the level of network resource. Given that a "truly" exogenous IV is very hard to find, as mentioned before, this study draws a great deal on Bayer and Ross (2009) to develop an IV to correct for selection bias. We know that to estimate Eq. (2) and Eq. (3) consistently, a standard assumption is that S_i and X_i are exogenous. However, due to the presence of selective friendship and other unobserved individual factors that enter ε_i, it is more plausible to assume that S_i is endogenous. I therefore assume that Cov $[X_i, \varepsilon_i] = 0$ and Cov $[S_i, \varepsilon_i] \neq 0$. Empirically, this means that network resource is endogenous due to the presence of selectivity, while the other observed individual attributes are assumed to be uncorrelated with any unobserved individual attributes. Bayer and Ross (2009) used these assumptions to construct a valid IV to take into account the selection problem. The approach consists of two steps, outlined below.

First, given that most individual attributes (e.g. educational attainment, gender, and age group) are discrete in nature, for the purpose of developing the IV I assign each respondent in the sample into a cell based on a set of his/her discrete

characteristics. That is, for each respondent *i*, I match to him/her the other respondents who are identical to himself or herself in terms of x_i. Second, for each respondent *i*, I average the network resource of the other members in the same cell and acquire the cell-averaged network resource (excluding the respondent himself/herself), which can be used to instrument for each respondent's own network resource. The justifications for this approach are:

1) The network resource s_i should be correlated with the IV because similar individuals, i.e. respondents in the common cell, make similar friendship choices. This, in fact, can be tested for in the first-stage regression of the IV estimator.

2) The constructed IV is uncorrelated with unobservables ε_i, because the forming process for the cell is entirely determined by x_i which are assumed to be exogenous under the standard assumption Cov $[x_i, \varepsilon_i]=0$. According to Bayer and Ross (2009), this IV approach amounts to using a non-parametric selection model to predict each individual's network resource, given their observable characteristics x_i.

Consequently, the IV model predicting labour market participation can be written as the following equation system：

$$y_i^* = \beta_0 + \beta_1 S_i + \beta_2 X_i + \varepsilon_i \ , \ y_i = 1[y_i^* > 0] \qquad (4)$$

$$S_i = \gamma_0 + \gamma_1 Z_i + \gamma_2 X_i + \xi \qquad (5)$$

where z_i is the average of network resource of observationally identical individuals (excluding himself or herself) and ξ is the random error term.

Likewise, the IV model predicting income is expressed by:

$$I_i = \beta_0 + \beta_1 S_i + \beta_2 X_i + \varepsilon_i \quad (6)$$

$$S_i = \gamma_0 + \gamma_1 Z_i + \gamma_2 X_i + \xi \quad (7)$$

Given Cov $[z_i, \varepsilon_i] = 0$, Cov $[\xi, \varepsilon_i] = 0$, and Cov $[z_i, S_i] \neq 0$, I regress S_i on Z_i in the first-stage regression (Eq. 5/7) and get $\hat{S}_i = \hat{\gamma}_0 + \hat{\gamma}_1 Z_i + \hat{\gamma}_2 X_i$. In the second-stage regression (Eq. 4/6), I use $\hat{S}_i$ as the replacement for S_i to acquire consistent estimates of β_1, since $\hat{S}_i$ must not be correlated with ε_i as long as the IV is valid. Notice that if the selection is driven by social homophily, the standard estimates of network effects would be upward-biased. However, if the selection is driven by social heterophily, the estimates would be downward-biased.

Before closing this section, an issue worth mentioning is the way in which one understands the implications of the critical working assumption that Cov $[X_i, \varepsilon_i] = 0$.[49] Regarding this, Bayer and Ross (2009:6) remark that it is reasonable to

49 For example, one may argue that the other controls such as individual educational attainment or Chinese Communist Party (CCP) membership may be correlated with some unobserved attributes, although the assumed exogeneity of controls is standard whenever we make a causal interpretation of the estimated correlation between the dependent variable and independent variable of interest.

impose this assumption if the aim of the analysis is to correct for the "additional bias that might arise due to selection" on the network resource variable. In the context of this study, the IV estimates of the network effects are more meaningful and preferred not because they are unbiased, but because they have taken into account self-selection as compared to the estimates obtained from any standard single-equation model. As mentioned in the introduction section, the aim of this chapter is to point out and correct for the selection problem rather than to precisely give the magnitude of the effects of educational network resource..

4. DATA AND VARIABLES

The data are from the 2003 CGSS, which was conducted jointly by the Survey Research Centre of the Hong Kong University of Science and Technology and the Sociology Department of Renmin University. The CGSS is an annual questionnaire-based survey of China's urban and rural households. Using a four-stage stratified sampling scheme, face-to-face interviews were conducted with 5,900 respondents in 125 counties/districts in mainland China. Notice that these respondents are urban residents in the sense that they lived and worked in the city in the surveying year, and their status in the Household Registration System were "Urban Residents". Thus rural-to-urban migrants were not interviewed. In the present study, I discarded respondents who were full-time students or farmers and those who were younger than 16 or older than 60 in the

surveying year. Due to missing values, the effective sample size is 4,026 for the employment analysis and 3,545 for the income analysis. Table 1 presents basic statistics regarding the respondents. The major explanatory variable, control variables and the developed IV variable are discussed below.

4.1 Network Resource Measure

Network resource here is measured with the average years of schooling of those included in the focal respondent's core discussion network. Using educational attainment to measure network resource is not unusual. Mouw (2003), for example, uses the average education of friends to measure network social capital and reports a positive association between it and the income of the respondent. In the 2003 CGSS, respondents were asked to name up to five people with whom they had discussed important matters within the last five months and to describe the educational attainment of the cited discussants. Network resource is thus calculated by averaging the years of schooling among one's named friends. As is shown in Table 1, the average network educational attainment among respondents is around 10.6 years. Notice that around one third of respondents cited five friends and the other two thirds named between one and four friends.

Using educational attainment to measure network resources does have its advantages. According to Mouw (2006), aside from self-selection, the other

source of endogeneity bias in estimating the role of network social capital is the simultaneity problem. That is, there might be mutual causality between the network resource variable and the outcome variable, leading to biased estimates. Since network resource is measured by average educational attainment among one's friends, the simultaneity problem can be comfortably sidestepped because it is reasonable to impose a temporal ordering on the network resource variable and the outcome variables to break the simultaneity: friends' educational attainments were achieved "earlier" than the respondent's "current" employment or income in the surveying year. So, there is no mutual causal relationship between the dependent variable and the independent variable.

One may also wonder whether or not these cited discussants are the alters where the network resources useful for job-seeking are indeed embedded. The interview protocol of the 2003 CGSS did not inform respondents about what was considered to be an "important matter". However, the survey did ask whether the discussion was about "a specific thing needing his/her influence and help" or about "emotional or social problems". The number of named people with whom the respondents discussed "emotional or social problems" accounted for merely 36.5 per cent of all the cited discussants, while those with whom respondents often talked about "a specific thing needing his/her influence and help" or talked about both accounted for 63.5 per cent. This implies that the core discussion network in China is a very important source of embedded resources for

instrumental actions.

Table 1: Selected descriptive statistics (N=4026)

Individual attributes	Mean	S.D.	Min	Max
Age	38.60	9.54	18	60
Age squared	1581.45	747.02	324	3600
Schooling years	10.87	3.40	0	20
Network educational attainment	10.60	3.04	0	20
Income （N=3545）	10281	14220	0	300000

	Freq.	per cent
Male	2028	50.37 %
Father is a CCP member	1109	27.55 %
Self-reported employed	2875	71.41 %
Father's supervising unit Rank 1	1579	39.22 %
Father's supervising unit Rank 2	285	7.98 %
Father's supervising unit Rank 3	1179	29.28 %
Father's supervising unit Rank 4	983	24.42 %
Region 1	1718	42.67 %
Region 2	727	18.06 %
Region 3	872	21.66 %
Region 4	709	17.61 %
Named 1 friend	724	17.98 %
Named 2 friends	780	19.37 %
Named 3 friends	754	18.73 %
Named 4 friends	400	9.94 %
Named 5 friends	1368	33.98 %

4.2 Labour Market Outcomes

The analysis in this chapter focuses on two outcomes: labour force participation and annual income. In China, due to the insufficient social insurance and welfare system, very few people choose not to participate in the labour market. In addition, in accordance with the social norm inherited from the centrally planned economy and rooted in traditional Confucianism, idleness has long been regarded as being shameful. In the 2003 CGSS, respondents were asked: "What were you mostly doing during the last three months: 1) full-time working; 2) part-time working; 3) temporary working; 4) farming; 5）unemployed, or 6) other" and: "How much was your income in 2002 (the total income including wages, bonus, allowance, stocks, and futures etc.)? " As is shown in Table 1, almost three quarters (71.41 per cent) of respondents in the sample were employed.

The second dependent variable is annual income. I have not used wage as the dependent variable because a person's social network not only influences his/her income from the job but also affects income from other channels, such as financial investments. Many respondents chose not to report their income in the survey. This means that only 3,545 observations can be included in the income analysis. Amongst them, the average income is around 10,281RMB (US $1,254). In addition, 530 out of the 3545 reported having an income of zero. I did not

exclude them because of two concerns. First, the Regulations on Minimum Subsistence Allowance for Urban Residents were not implemented by the Chinese government at a country-wide level until 2003. That is, there were some urban residents who had no pension or unemployment insurance in 2002. Second, it is the annual total income that is of interest to this study, and income from financial investments such as stocks and futures can be negative. That being said, however, in the robustness check I will use the working sample without zero income to see whether the major results are sensitive to it. Also during the robustness check, I will drop the income outliers. Finally, it should be noted that, when fitting the model, the annual income appears in logarithmic form. Regarding respondents with zero income, I recoded their income from zero to 1 RMB and get the logarithmic value zero.

4.3 Other Controls

As can be seen from the table above, a set of variables measuring different individual characteristics is included in the model to control for other factors influencing labour market participation and income. They are the individual's age, age squared, sex, schooling years, father's membership of the CCP, the rank of the supervising unit of the respondent's father when the respondent was 18 years old and region dummies. One should note that in the 2003 CGSS data, only 220 of 4,026 respondents are not Han Chinese. As a result, I did not include

ethnic dummies. In addition, I did not include individuals' CCP membership as most people join the CCP in their work units.

I control for fathers' membership of the CCP and the rank of fathers' supervising unit because political status is important to job outcomes in China (Li and Walder 2001; Walder 2003). Although the importance of the role of party membership is believed to have declined since the ruling party introduced market reforms in the late 1980s, it is still dominant. The rank of the supervising unit also measures the political capital of the respondent's family, because the administrative rank within the state hierarchy constitutes the most significant component of social stratification and social mobility in China under state socialism (Lin and Bian 1991; Bian 1994). As can be seen in Table 1, the father's membership of the CCP is measured with a binary dummy, while the rank of the father's supervising unit is measured on four levels: without supervising unit (Rank 1); city/county/district government (Rank 2); provincial government (Rank 3); and central ministry and above (Rank 4)[50].

Given regional differences in socioeconomic development, regional dummies are incorporated into the model. As face-to-face interviews were conducted in the houses of urban residents, the regional dummies are based on the place of residence.[51] According to per capita GDP from 2002, the surveyed provinces are

50 Lin (2004) also uses the rank of the supervising unit to proxy social resources.

51 Given the huge geographic scale, the "region" of the respondent is based on both place of residence

classified into four groups:

- Region 1 with a GDP of more than US$10,000: Beijing, Shanghai, Tianjin, Jiangsu, Zhejiang, Guangdong, Shandong, Fujian and Liaoning.
- Region 2 with a GDP between US$7,000 and US$10,000: Xinjiang, Hebei, Hainan, Jilin, Hubei and Hunan.
- Region 3 with a GDP between US$5,000 and US$7,000: Henan, Shanxi, Heilongjiang, Ningxia, Anhui, Chongqing, Qinghai and Sichuan.
- Region 4 with a GDP of less than US$5,000: Shaanxi, Yunnan, Jiangxi, Guangxi, Gansu, Inner Mongolia and Guizhou.

Importantly, I use the region dummy rather than provincial dummy as controls to make sure the as many respondents can be allocated to a certain cell as possible. The geographic distribution of the regions is illustrated in Figure 1. Notice that other areas not surveyed in the 2003 CGSS are labeled as Region 5.

There are, of course, many more factors influencing employment and income, and more detailed analysis can be conducted, but here I focus on major predictors and control for the factors that are provided by the data. Another concern regarding not including too many controls is because, if more were included, the sizes of the constructed cells of observationally identical individuals

and work.

would be too small.

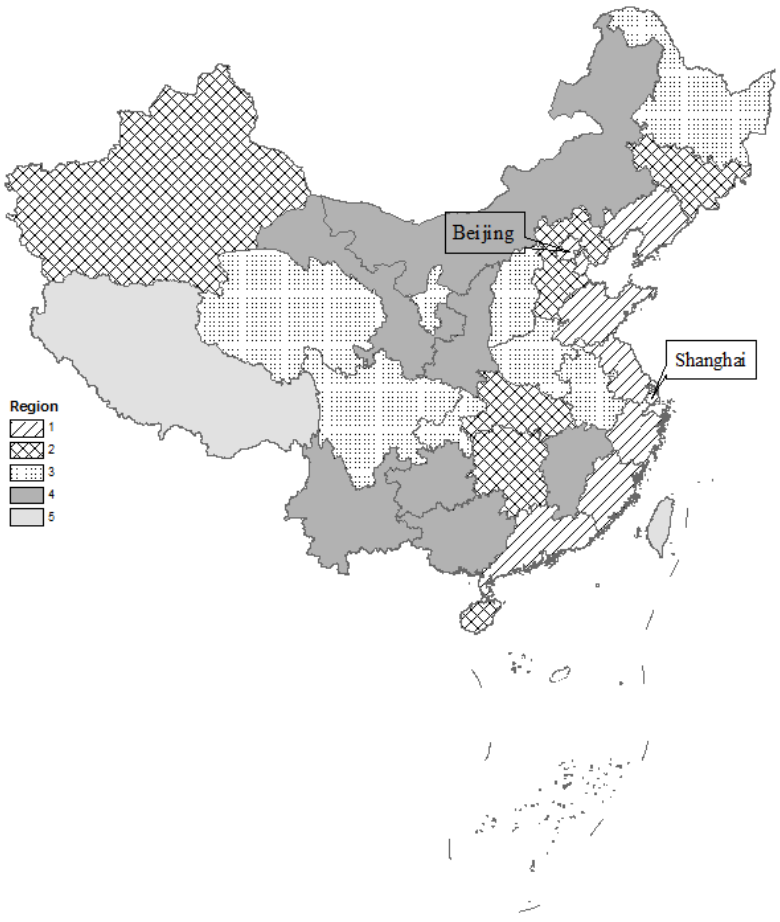


Figure 1: Regions of economic geography

4.4 Instrumental Variable

For the purpose of constructing the IV, I assign each individual to a cell based on the set of his/her characteristics (controls) described in Table 1. I use three categories for age (from 16 to 30 years old, from 30 to 40 years old, and from 40

to 60 years old); two for sex (male and female); three for schooling (less than 13 years, from 13 to 16 years, and more than 17 years); two for father's CCP membership (being a CCP member and not being a CCP member); four for the administrative rank of the father's supervising unit (without rank, city/county/district level, province level, and central ministry level); and four for regions.

Based on all the possible permutations of these categories, I arrive at 576 cells for 4,026 respondents in the 2003 CGSS data. The size of cells ranges from 0 to 116, and the Per Capita cell size is 33. Individuals belonging to the same cell as the foci individual are used to calculate the average of educational network resource in the cell (excepting the foci individual himself or herself) and this average forms the IV. One should note that there were 52 respondents who could not be matched to any other individuals, so these 52 respondents were dropped from the IV analysis. To reduce the potential measurement error further, in the robustness check I will also drop small cells. That is, I will examine how the results are affected when excluding individuals who are matched to cells smaller than 3, 4 and 10 respectively.

5. RESULTS FROM SINGLE-EQUATION MODELS

Table 2：Single-Equation Models of Employment and Log Income

	Employment as Dependent	Income as Dependent
	Probit Model	OLS Model
Independents	**Model 1a**	**Model 2a**
Educational network resource	.065***	.075***
	(.00904)	(.0173)
Age	.045***	.15***
	(.0175)	(.0378)
Age squared	-.00069***	-.002***
	(.000022)	(.00047)
Schooling Years	.089***	.10***
	(.00885)	(.0168)
Father's CCP Membership	.19***	.17*
	(.0557)	(.0927)
Father's Supervising Unit Rank	.19***	.12
	(.0603)	(.112)
Father's Supervising Unit Rank	.21***	.011
	(.0574)	(.106)
Father's Supervising Unit Rank	.072	.067
	(.0982)	(.151)
Female	-.58***	-.93***
	(.0458)	(.0853)
Employment	--	4.24***
	--	(.141)
Region 2	-.002	-.49***
	.0650	(.106)
Region 3	-.15**	-.52***
	(.0591)	(.117)
Region 4	-.17***	-.72***
	(.0632)	(.122)
Constant	-1.83***	-.95
	(.347)	(.745)
N	4026	3545
R-square	0.14	0.46

Note: i) Standard errors are in parentheses and robust to heteroscedasticy;
ii) * p < =0.1 ** p <= 0.05, *** =p < 0.01
iii) Males, being unemployed, Father's Rank 1, and Region 1 as reference groups.

In this section, I treat network resource as an exogenous variable and, therefore, estimate the single-equation probit and OLS models. The major results, estimated coefficients and relevant p-values are presented below in Table 2. Model 1a predicts the probability of being employed, while model 2a predicts annual income. As is shown in Table 2, the network resource measured with average educational attainments of one's core discussion network significantly improves both the probability of being employed and the level of annual income. For example, according to the results of model 1a, one more year of network average schooling is associated with a two percentage higher probability of being employed.[52] According to model 2a, one more year of network average schooling is related to an increase in annual income by 7.5 percent.

It is clear that both models point to the important role of educational attainment. However, the ascriptive features are also important. For example, father's political status and the rank of his supervising unit also play important roles in getting a job and attaining income. This pattern reflects the social transition that China is undergoing: the effect of educational attainment reflects the market transition, while the influence of family background in terms of political loyalty and hierarchy reflects the extent to which social resources are still controlled by the government.[53] An interesting finding is in the effect of father's supervising unit rank. I found that the employment effect of Rank 4 (central ministry or

52 The straight coefficient 0.065 denotes a marginal effect of 0.02.

53 The debates surrounding China's social transition are covered in Bian and Ang (1997), Nee (1989, 1991) and Cao and Nee (2000).

above) is even higher than Rank 3 (provincial). This might be explained by the subtle gaming-relationship between the central and local governments: local governments are supposed to be controlled directly by Beijing but in reality they are more often controlled by local bosses, especially at the provincial level. A Chinese proverb vividly reflects this situation: the mighty dragon is no match for the native serpent. Finally, with all other factors held constant, age influences employment and income in a non-linear way. As expected, the regional dummies are significant, reflecting the huge regional differences in levels of development in China.

6. RESULTS FROM IV MODELS

First of all, I examine the quality of the IV. As can be seen at the bottom of Table 3, the F-statistics of the first-stage regression are 38.3 for the employment IV estimator and 29.8 for the income IV estimator. The rule of thumb (see Cameron and Trivedi 2005) is that the first-stage F-statistics should be larger than 10 for models with single endogenous regressors(the estimates of the first-statge regression of Model 1c and Model 2c can be seen in Appendix 1 and Appendix 2). Furthermore, Stock and Yogo's (2005) weak IV critical value is 16.38 for one endogenous regressor and one IV. Consequently, the two F-tests comfortably pass the underidentification and weak-identification test, indicating that the endogenous variable is strongly associated with the IV. More importantly, the

Hausman endogeneity test (for the OLS model) and the Wald endogeneity test (for the probit model) indicate that one can reject the null hypotheses that there are no systematic differences between the estimates from the OLS/probit and IV models. Since the IV estimator always yields consistent estimates given a valid IV, its coefficient estimates are more meaningful.

To compare the IV estimates with estimates from standard single-equation models, I dropped observations that have missing values on the IV. This reduced the sample size of the employment model from 4026 (model 1a in Table 2) to 3974 (model 1b in Table 3), and the sample size of the income model from 3545 (model 2a in Table 2) to 3496 (model 2b in Table 3). Notice that the results obtained from single-equation estimates in Table 3 are very similar to those in Table 2, suggesting that slightly smaller samples (52 observations dropped in employment analysis and 49 observations dropped in income analysis) do not affect the estimated results in a significant way. A comparison of the estimates from the IV and single-equation models demonstrates that, although network effects remain significant, accounting for self-selection driven by individual unobservables has substantially increased the estimated effects on employment from 0.066 to 0.26 and the effects on income from 0.075 to 0.73. As for the other controls, most of the results are roughly similar with or without corrections for self-selection. The two exceptions are education and father's CCP membership when I include the IV.

Table 3：IV Models on Employment and Log Income

	Employment as Dependent		Income as Dependent	
Independents	**Model 1b (Probit)**	**Model 1c (IV-Probit)**	**Model 2b (OLS)**	**Model 2c (IV)**
Educational network resource	.066***	.26***	.075***	.73***
	(.0091)	(.0643)	(.0175)	(.212)
Age	.044***	.060***	.15***	.22***
	(.0176)	(.0169)	(.038)	(.0489)
Age squared	-.00068***	-.0009***	-.0016***	-.0003***
	(.00022)	(.00002)	(.0004)	(.00063)
Schooling years	.087***	.026	.104***	.23**
	(.00893)	(.0453)	(.0172)	(.112)
Father's CCP Membership	.18***	.068	.18*	-.069
	(.0561)	(.070)	(.0946)	(.140)
Father's Supervising Unit Rank	.19***	.20***	.129	.22
	(0.604)	(.0571)	(.113)	(.135)
Father's Supervising Unit Rank	.20***	.22***	.017	.12
	(.0573)	(.0540)	(.107)	(.129)
Father's Supervising Unit Rank	.076	.124	.018	.14
	(.1002)	(.0901)	(.156)	(.183)
Female	-.589***	-.54***	-.93***	-1.08***
	(.046)	(.0631)	(.0866)	(.114)
Employment	--	--	4.23***	3.79***
	--	--	(.142)	(.213)
Region 2	-.011	-.029	-.49***	-.42***
	(.0653)	(.0625)	(.108)	(.141)
Region 3	-.149**	-.12**	-.52***	-.45***
	(.0593)	(.0637)	(.118)	(.129)
Region 4	-.162**	-.12**	-.71***	-.77***
	(.0634)	(.0578)	(.123)	(.147)
Constant	-1.79	-2.88	-.902	-4.89
	(.348)	(.435)	(.750)	(1.54)
First-stage F-statistic	--	38.3	--	29.8
Hausman test of endogeneity	--	--	--	Chi(1)=14.2 p=0.0002
Wald test of endogeneity	--	Chi2(1)= 5.69 p=0.017		
N	3974	3974	3496	3496

Note: i) Standard errors are in parentheses and robust to heteroscedasticy; ii) * p < =0.1 ** p <= 0.05, *** =p < 0.01;
iii) Males, being unemployed, Father's rank 1, and Region 1 as reference groups.

An important result, therefore, is that the error terms in the single-equation outcome models must be negatively correlated with the error terms in the selection models. This is the case because the single-equation models underestimated the network effects by mistakenly offsetting part of the impact of the unobservable individual characteristics, i.e. selection factors such as ability or IQ. That is, unobserved selection factors are associated with both labour market outcomes and average education of friends, but in opposite directions. The findings therefore underscore the significant role of network resources and suggest the presence of heterophilous selection in the formation of core discussion networks.

Table 4 presents the results of a robustness check to see if the IV estimates of network effects are sensitive to alternative measures of the major explanatory variables or different sub-samples. First, I use the average network educational attainment weighted by the tie strength (i.e. the order in which the respondent cited his or her friends) as the explanatory variable. The tie strength was coded as 5 for the first friend, 4 for the second, 3 for the third, 2 for the fourth and 1 for the last friend. Second, I incorporate provincial dummies into the model.[54] Third, as discussed above, I drop cells smaller than 3, 4 and 10 so as to reduce the effects of possible measurement errors. These restrictions reduce the sample size by 132, 258 and 1034 observations respectively. Fourth, I drop respondents who

[54] Notice that the cells are still developed by regions.

reported zero income in 2002, which reduced the sample by 530 individuals. Fifth, considering the potential outliers in terms of income, I delete 151 observations in the top 3 per cent. The results of the five tests are all presented in Table 4.

Table 4：Robustness Check for IV Models

Employment as Dependent		Income as Dependent	
Estimated effects		**Estimated effects**	
Model 1c	.26***	**Model 2c**	.73***
Model 1d	.25***	**Model 2d**	.74***
Model 1e	.30***	**Model 2e**	.58**
Model 1f	.26***	**Model 2f**	.68***
Model 1g	.24***	**Model 2g**	.68***
Model 1h	.30***	**Model 2h**	.67**
		Model 2i	.54**
		Model 2j	.50***

Note: (1) Models 1d and 2d use the weighted mean of education to measure network resource; (2) Models 1e and 2e use provincial dummies rather than regional dummies; (3) Models 1f and 2f are fitted on the sample excluding cells smaller than 3; models 1g and 2g are fitted on the sample excluding cells smaller than 4; models 1h and 2h are fitted on the sample excluding cells smaller than 10; (4) Model 2i is fitted on the sample without zero income; (5) Model 2j is fitted on the sample excluding the top 3 per cent of income.

For the purpose of comparison, the first row of Table 4 repeats the basic IV estimates in Table 3 (i.e. models 1c and 2c). As can be seen in Table 4, all estimates are similar to those in the first row and roughly equal to 0.3 for the employment model and 0.7 for the income model. The main exceptions are income model estimates in models 2i and 2j, which excluded income outliners. A likely reason is that the lowest-income and highest-income groups benefit less

from core discussion networks as compared to middle-income groups.

7. INTERPRETATION AND DISCUSSION

As mentioned above, a likely reason for the IV estimates of the network effect being larger than those obtained from single-equation models is that individuals with poor unobservables in terms of employment or income might compensate for those unobservables (unobserved to sociologists) by selecting friends with a higher level of educational attainment. Although this looks counterintuitive, similar findings have been found in neighbourhood studies. Based on the data from the 1990 Decennial Census for the Boston MSA, Bayer and Ross (2009) find that the IV estimates of the effects of geographic proximity to jobs, neighbourhood poverty rates and average neighbourhood education are substantially larger than the correlations identified using OLS. They thus argue that, "the finding that OLS estimates of neighbourhood effects are biased downwards is consistent with the hypothesis that individuals with poor unobservables in terms of labour market outcomes compensate for these unobservables by sorting into locations with better employment prospects" (Bayer and Ross 2009: 25).

How is heterophilous selection possible in a world where we believe that people

are more likely to socialise with those who are similar to them? Notice that, although social homophily is pervasive in the formation of social ties, it is not necessarily the case that social homophily is the exclusive selection principle. In fact, some empirical studies show that homophilous selection is not as pervasive as one may expect. In a comprehensive review of research on homophily, McPherson, Smith-Lovin and Cook (2001) concede that, although many demographic factors such as race, gender, age and religion give rise to stark divisions in the tie-formation process, other individual attributes including education, occupation and network position do not necessarily lead to homophily-based selection. For example, Verbrugge (1977) reports that occupational similarity is not as important for the selection of best friends as it is for second- and third-ranking friends. Similarly, Louch (2000) finds that educational similarity is less significant among confidants in comparison to religious or racial similarity.

These studies imply that the presence of social homophily along a set of dimensions does not preclude the possibility that individuals also choose some friends dissimilar to themselves along some other dimensions. As Lin (2008) argues, it is problematic to assume that homophilous and heterophilous networks are an either-or choice. As a matter of fact, the motivations behind the selection of friends and thereby network resources can be very complex. In general, heterophilous selection can be expected for four main reasons.

First, heterophilous selection is often beneficial for instrumental actions given the inequality of social resources. At the individual level, Briggs (1998) distinguishes between two types of social capital residing in relational networks: social support as aids for "getting by" and social leverage as aids for "getting ahead". Lin (2008) refers to the same categories as homophilous (or bonding) social capital with expressive utility, and heterophilous (or bridging) social capital with instrumental utility. In general, homophilous selection takes precedence over heterophilous selection in ordinary and normative forms of social relations. However, when the formation of social ties is guided or influenced by instrumental goals, heterophilous selection may arise because heterophilous resources embedded in social networks are more useful to mitigate socioeconomic inequality and make more diverse resources present in the focal network (Briggs 1998, 2002; Wellman and Gulia 1999; Dominguez and Watkins 2003).

Second, heterophilous selection makes it possible to form intimate relationships because heterophilous ties can be mutually beneficial for actors positioned differently on the socioeconomic ladder. For example, Robins and Boldero (2003) argue that relational discrepancy between two individuals helps to form and maintain seemingly unsymmetrical relations for two reasons. On the one hand, the lower-positioned individual in the unsymmetrical relationship may be

prepared to accord the other a "leadership" role and to adopt a "follower" role to ameliorate the adverse emotional reactions derived from the dissimilarity. On the other hand, the better-positioned individual may consider himself/herself to be "better" than the other, thereby gaining a pleasant reminder of their perceived "superiority". Thus, the advantaged individual may be motivated to maintain the relationship, even though he/she will not gain other resources from the social tie. Empirical evidence can be found in Robins and Boldero (2003) and Domingues and Watkins (2003).

Third, heterophilous selection can arise because of meeting opportunities that are influenced by certain social contexts. Although there is a large body of classical literature focusing on individuals' preference for certain types of social ties (e.g. Lazarsfeld and Merton 1954; Newcomb, 1961; Laumann, 1966), Blau (1977) argues that individual preferences are not the only factor determining the composition of ego-centric networks, that is to say that the formation of social ties can hardly be a merely personal decision. Instead, the process of relation-formation also depends on people's meeting opportunities which are structured by the demographic and socioeconomic composition of their social contexts.[55] For example, in a country with high inequality of social resources, preference for similarity may not be that important for selecting friends, because those who are positioned in the lower level of the resource ladder have to build

55 For a recent empirical study on this, refer to Mollenhorst, Völker and Flap (2008).

up more social ties toward the top of the ladder. Another example is how different group sizes may also lead to cross-group ties. The reason is straightforward: since each cross-group friendship must involve a member from each group, smaller groups must have more cross-group friendships on a per-capita basis. If we believe that social structure of resources, such as wealth and occupational prestige, is shaped as a pyramid, better-positioned individuals must have more heterophilous relationships than those who are at the bottom of the hierarchy.

Fourth, heterophilous selection is likely to be found on 'unfruitful' resources such as education in China under social transition. Notice that under state socialism returns to education are generally not as important as returns to positional power or type of organizations. Although Zhou (2000) reports increased returns to education in the reform era in China as compared to those in the pre-reform era, he reveals strong institutional persistence as returns to party membership and cadre status remain pivotal. In this context, more able or advantaged Chinese tend to socialise with those who possess more 'fruitful' resources. For those with poor ability or unobserved disadvantages, what they can do to "get ahead" is to turn to people with more 'unfruitful' resources such as educational social capital.

If the above interpretation of 'unfruitful' resource makes sense, one could expect that heterophilous selection to educational social capital is unlikely to be found in a labour

market where returns to education are substantive. To test for this, I applied the same IV approach to the data from the 1987 Urban Poverty and Family Life Study (UPFLS) [56]. Although not reported here, as expected, the estimates show no evidence of presence of heterophilous selection to educational social capital[57].

Again, we should bear in minds that in the context of this study, the IV estimates of the network effects are meaningful not because it is more precise but because they have taken into account self-selection which is the major source of endogeneity in social capital studies. That is exactly the reason why the estimates from OLS and probit models are also presented and compared. Although the coefficient magnitudes of IV models look implausibly large on a substantive level, the value of IV estimates is that they shed more lights onto the direction of how selectivity influence casual inference in estimating the role of embedded resources. Obviously, the IV estimates reported in this chapter are not definitive. On the one hand, we should not overreach the findings; on the other, we do not need to interpret them too much on a substantive level.

56 The 1987 UPFLS was involved 2,490 respondents aged 18-47 and living in poverty-stricken areas in Chicago. It asked respondents to name up to three best friends and their demographic information, such as schooling years, sex, and ethnicity. Using the OLS model, Mouw (2003) has analysed the 1987 UPFL data and reported a significant correlation between the average network education and the individual income. Yet, he aptly points out that the estimation is still open to the endogeneity problem.

57 In this test the strength of the correlation between IV and educational social capital is not that strong. This can probably be explained by two aspects. First, the method of developing IV requires a large sample, yet the sample size of UPFLS is small. Second, similar people in the US may not have similar "criteria" to select friends. In this regard, this IV method is to some extent ad hoc, and further studies to assess the argument of 'unfruitful' resources are called for.

8. CONCLUSION

People select friends or networks and thereby also the embedded resource available to them. This raises the possibility that the association between network resource variables and individual outcomes found in numerous studies is not a causal relationship, but actually reflects a correlation between other unobserved factors and network resources. The sociological literature on network effects and social capital has started recently to pay systematic attention to the selection problem (see Mouw 2003, 2006). To address this issue, this chapter uses an innovative IV approach to correct for the estimation bias. The empirical results reported here are robust across a wide variety of specification and robustness checks, and they indicate that standard estimates of the effects of educational network resource on employment and income are downwardly biased if one does not take into account self-selection.

This study therefore reveals that people with poor unobservables tend to adopt a compensation strategy when they socialise with others. It thus contributes to the existing literature on social networks by suggesting the presence of heterophilous selection in network formation. To the author's best knowledge, this issue has not been previously empirically explored in the sociological literature. What has been found here echoes Lin's (2008) important points that we cannot assume that a high degree of homophily over some characteristics

implies a low degree of heterophily across all characteristics.

Instrumental goals, asymmetric relationships, meeting opportunities and 'unfruitful' resources that are shaped by social contexts can all lead to heterophilous selection in the formation of friendships. Given the inequality of social resources, this class of selection is especially possible for those who are poorly positioned on the ladder of social resources. As pointed out at the outset of the paper, highlighting the presence of heterophilous selection in relationship formation should not be construed as doubting the weight of the principle of social homophily. Rather, the finding here complements Mouw (2003) and echoes his (2006) call to address potential endogeneity problems so as to identify the causal effects of social network.

Around 2,000 years ago, Cicero (106 B.C.–43 B.C.) concluded that:

> I gather that friendship springs from a natural impulse, rather than a wish for help: from an inclination of the heart, combined with a certain instinctive feeling of love, rather than from a deliberate calculation of the material advantage it was likely to confer.

However, this study suggests that friendship, at least in modern Chinese society, is a much more complex and multifaceted process than this.

Appendix 1 Estimates of the First-Stage Regression of Model 1c

Educational network resource as Dependent (N=3974)	
Independent Variables	
Instrumental Variable	.20***
	(.0324)
Age	-.087***
	(.0282)
Age squared	.0013***
	(.00036)
Schooling years	.46***
	(.019)
Father's CCP Membership	.307***
	(.087)
Father's Supervising Unit Rank 4	.090
	(.101)
Father's Supervising Unit Rank 3	.078
	(.098)
Father's Supervising Unit Rank 2	.14
	(.162)
Female	-.14*
	(.078)
Region 2	-.084
	(.106)
Region 3	.006*
	(.099)
Region 4	.12
	(.112)
Constant	4.71***
	(.591)

Note: i) Standard errors are in parentheses and robust to heteroscedasticy; ii) * p < =0.1 ** p <= 0.05, *** =p < 0.01;
iii) Males, being unemployed, Father's rank 1, and Region 1 as reference groups.

Appendix 2 Estimates of the First-Stage Regression of Model 2c

Educational network resource as Dependent (N=3974)	
Independent Variables	
Instrumental Variable	.18***
	(.0346)
Age	-.083**
	(.0303)
Age squared	.0012**
	(.00039)
Schooling years	.45***
	(.0212)
Father's CCP Membership	.26**
	(.0922)
Father's Supervising Unit Rank 4	.084
	(.108)
Father's Supervising Unit Rank 3	.084
	(.103)
Father's Supervising Unit Rank 2	.077
	(.167)
Female	-.24**
	(.0824)
Region 2	-.12
	(.112)
Region 3	-.017
	(.104)
Region 4	.099
	(.118)
Employment	.67***
	(.0983)
Constant	4.55***
	(.631)

Note: i) Standard errors are in parentheses and robust to heteroscedasticy; ii) * p < =0.1 ** p <= 0.05, *** =p < 0.01; iii) Males, being unemployed, Father's rank 1, and Region 1 as reference groups.

Chapter 5

The Effects of Using Contacts under State Socialism: Evidence from the First Difference Analysis

1. INTRODUCTION

A large body of sociological literature reveals that some people find jobs through informal channels, that is, through help from contacts or network members, while others find their jobs through formal channels such as agencies or job advertisements (Granovetter 1973; Lin, Ensel, and Vaughn 1981; De Graaf and Flap 1988; Marsden and Hurlbert 1988; Wegener 1991; Bian 1997; Völker and Flap 1999).[58] It is therefore of interest to see whether using contacts to find a job pays off in comparison to using other channels. One of Granovetter's (1973, 1974) major hypotheses is that the use of contacts allows individuals to gain advantages in the labour market by providing more and better information about job entries. Likewise, social resource theory (Lin 1999) also implies the

58 Economic literature has also found similar patterns. For example, in a recent study based on a European Community Household Panel (ECHP1994-1999), Pellizzari (2009) finds that informal contacts are widely used by both firms and workers in most European countries. The estimated probability of finding a job using contacts is 45 per cent for Spain, 38.5 per cent for Portugal, 41 per cent for Greece, 37.9 per cent for Italy, 34.4 per cent for France, 32.2 per cent for Germany, 23 per cent for the UK, 13.3 per cent for Finland, and 17.7 per cent for Denmark.

importance of social contacts, since social resources such as job-related information or influence are embedded in social ties.[59]

Although there is limited sociological analysis focusing on the general benefits of using contacts in comparison with other job-seeking methods, there has been a research tradition of such analysis in studies estimating the effects of the characteristics of contacts (e.g. the occupational prestige of a given job-helper). As will be shown in the literature review section, the relevant findings are mixed. Although some studies report the benefits of using contacts, most of the existing findings indicate that no statistically significant effects are produced by using contacts on job outcomes such as wages, occupational prestige, unemployment duration, job satisfaction, and job tenure (for extensive reviews see Lin 1999 and Marsden and Gorman 2001).

However, the estimated effects of using contacts in previous sociological studies are called into question by a potential endogeneity problem. As pointed out by Mouw (2003), most social capital studies in sociology estimating the direct role of using contacts treat the explanatory variable as exogenous. That is, using or not using contacts is regarded as a random choice. However, this assumption of exogeneity can hardly hold, as people may s**e**lect a certain method and this class

59 Ioannides and Datcher-Loury (2004) and Calvó-Armengol and Ioannides (2005) distinguish two mechanisms (from the employee's perspective and from the employer's perspective) through which social networks influence the functioning of labour markets. Here I focus on the employee's use of contacts.

of selection is often driven by unobserved individual characteristics which might also influence individual labour market outcomes. How then to get around the selection problem? Provided that the individual heterogeneity is time-invariant, fitting a fixed-effects model or first-difference model on panel data could alleviate the selection bias by eliminating the time-fixed unobservables from the model. When it comes to relevant empirical analysis in sociology, however, there are very few studies exploiting panel data except for those by Mouw (2002, 2003) focusing on US labour markets.

The other issue is that, although Mouw (2002, 2003) to a large extent confirms previous findings that using contacts has little direct role in a free market economy, so far little attention has been paid to alternative social contexts such as those experienced in countries under state socialism[60]. Völker and Flap (1999) analyse the data collected in the GDR and also report no evidence of a direct role being played by using contacts. However, as mentioned above, their cross-sectional method does not adequately deal with the potential selection bias.[61] To build upon the studies of Mouw (2002, 2003) and Völker and Flap (1999) then, this chapter looks into two unique data sets from countries governed by state socialism.

60 State socialism is contrasted with forms of socialism that advocate direct self-management, adhocracy and direct cooperative ownership of the means of production. See the definition and features of state socialism in Zhou and Suhomlinova (2000)

61 Although Völker and Flap (1999) collected two-period panel data (first job and job in 1989), they actually fitted two OLS models for the two time periods.

More specifically, I replicate and extend Völker and Flap's study (1999) on the GDR data. Using the same models, I also analyse the data from 2002 CHIPS. Both the GDR and the Chinese results are consistent with previous findings, showing no consistent evidence supporting the benefits of the use of contacts by itself as compared to the use of other channels to find jobs. However, the GDR study reveals that the effect of using contacts on status attainment is upwardly biased if one does not take into account potential selection, while the Chinese study shows that the standard OLS estimate of wage effect is downwardly biased. This suggests that more disadvantaged workers with respect to unobserved social background, achievement and ability in China tend to use contacts to find jobs, while in the GDR this is not the case. It contributes to the existing literature not only methodologically, but also by revealing that there are different mechanisms through which individuals choose job-seeking strategies in different countries.

The rest of the paper is structured as follows. Section 2 provides a literature review of recent studies into the effects of using contacts, emphasising that the presence of selection behaviour may plague causal inference by biasing the estimated effects. Section 3 outlines the first-difference method and its usefulness for this type of research. Section 4 analyses the GDR data and Section 5 describes and analyses the Chinese data. Section 6 offers an interpretation of the findings, paying close attention to the different selection patterns between Chinese and German individuals. Finally, Section 7 draws conclusions from the

analysis.

2. LITERATURE REVIEW

The pioneering work by Granovetter (1973) distinguishes three channels through which people get job leads: personal contacts (e.g. relatives, workmates, friends), formal channels (e.g. job announcements, hiring offices) and direct contact with employers. In this study, I focus on the effects of personal contacts. As mentioned above, there is very little previous sociological research exclusively addressing the direct impact of using personal contacts on labour market outcomes. Rather, the research tradition of many social capital studies probing the effects of social capital (e.g. the effect of contacts' prestige) is to examine the direct role of using contacts in the first place. In addition, some recent empirical econometrics studies address this issue, exploiting useful identification strategies to deal with the potential selection problem. Despite being born of different disciplines, the methods used in these papers to help identify models are useful for sociological analysis.

Although there is a small amount of early literature empirically supporting the positive association between using contacts and job outcomes, most studies have not reported significant benefits (see in Chapter 2). And, with the exception of Völker and Flap (1999), most studies examine the role of using contacts or

informal networks in the US and Europe, i.e. within market economies. Little attention has been paid to countries under state socialism. More importantly, most early studies treated the explanatory variable as exogenous and simply fitted OLS models onto cross-section data.

Among relatively recent research, four studies exploit fixed effects models to alleviate the selection bias. Using longitudinal data from the National Longitudinal Study of Youth (NLSY), Mouw (2002) compares cross-sectional and panel results on the relationship between the use of contacts and wages among black people in the US. Although the cross-sectional data demonstrate a significant and negative effect from using contacts, the results obtained from the fixed-effects model on the longitudinal data suggest no role whatsoever for contacts in the process. Mouw (2002) argues that this might be due to the lower levels of opportunity among workers who use contacts rather than the effect of contacts per se. That is to say, the opportunity of job offers is the omitted variable from which the selection has arisen. As a result, the OLS model underestimates the true effect of using contacts by mistakenly offsetting the role of the opportunities of a job offer. Next, Mouw (2003) fits a fixed-effects model on panel data (NLSY 1994–1998) and finds that there is no significant association between using contacts and wages in the US in the 1990s. This is consistent with his early findings among black Americans.

Based on the NLSY data from another time period (1981-1982), Kugler (2003) generates evidence using a first-difference model showing that the positive and significant OLS and first-difference estimates of the effects of using employee referrals become insignificant once employment sector is controlled for. Using panel data, Pellizzari (2009) conducts a comparative study of European countries. He finds that informal search channels lead to significantly better-paying jobs in Austria, Belgium and the Netherlands, while the opposite is true in Greece, Italy, Portugal, and the UK. In other EU countries – as well as in the US – no significant wage difference is observed. Furthermore, an even larger variation emerges from the cross-industry analysis. He then brings to the fore a further explanation of the wage effect of referrals and the variance across countries. A negative wage effect would be typical of labour markets where formal recruitment methods are more common. In this case, informal channels are generally considered a cheap selection device and are therefore not as effective.

Aside from using fixed effects model or first difference model, IV approach has also been employed. Bentolia, Michelacci and Suarez (2003) use the number of older siblings and the US state where the worker lived at 16 years of age to instrument using contacts or not. With IV estimates, they contend that, because of the mismatch between a worker's occupational comparative advantage and his/her actual productivity, the use of contacts actually leads to lower wages. In addition to exploiting advanced identification strategies, research undertaken

after 2002 also controls for more variables to alleviate the potential endogeneity problem. Especially, more attention has been paid to the role of the types of contacts and occupations of respondents which could be omitted factors leading to selection bias. For example, Mosca and Pastore (2008) find that the role of using contacts is not uniform across different sectors. Loury (2006) reports a negative wage premium for men relying on female contacts, a positive wage premium for men relying on male contacts, a positive wage premium for women relying on male contacts, and an insignificant effect for women relying on female contacts.

Based on the above literature, we cannot arrive at any conclusion that the effects of using contacts in previous sociological studies are in general overstated or understated simply because their findings are mixed. Theoretically, to find how the estimates are biased, one should fully examine how the omitted individual factor on which selectivity arises is correlated with the individual outcomes of interest and the using of contacts. That is, the direction of the bias depends on the relationship between the omitted individual factor and the choice of job-seeking method. For example, when people who are disadvantaged in the labour market (e.g. those who have less ability, those with fewer job opportunities, or those in lower-waged industries) tend to use contacts, and these factors are not controlled for, the standard estimate of the direct effect of using contacts would be downwardly biased because it mistakenly offsets the

role of those omitted variables on which selection may arise. In contrast, if those who have advantages in the labour market choose to use contacts, the standard estimate of the general benefit of using contacts would be upwardly biased, as the estimate must represent a combination of the "real" role of using contacts and the effect of those omitted variables. Since selection behaviour may be different in countries even under similar political contexts such as state socialism, a comparative analysis is helpful as it could shed more lights on our understanding about how people select job-search in different ways.

3. THE FIRST-DIFFERENCE MODEL

The basic specification of the model includes the major explanatory variable of interest (which is a dummy variable coded as 1 if respondents used contacts to find themselves a job and 0 otherwise), the observed individual characteristics as controls (e.g. age, age-squared, sex, educational attainment, etc.), and unobserved individual heterogeneity. Taking, for example, wages as the dependent variable, to estimate the effect of using contacts on wages, one can start with a standard OLS model which can be expressed by:

$$W_i = \beta_0 + \beta_1 C_i + \beta_2 X_i + \alpha_i + \epsilon_i \quad (1)$$

where W_i is the wage of the individual *i*. C_i denotes the use of contacts, X_i denotes a vector of individual factors, α_i is the individual unobservables influencing wages, and ϵ_i is the random error. Notice that C_i = 1 if contacts are

used. In Eq. (1), β_1 is the difference in the constant absolute value of wages for those who used contacts to find jobs rather than non-users. This method, as mentioned before, is confronted by a major difficulty: it considers the uses of contacts as exogenous or random rather than selective, and leads to selection bias.

There is a set of identification strategies available to deal with the endogeneity problems. This chapter exploits the informative two-time-period data sets from the GDR and China, and estimates the first-difference models. Notice that two-time-period data sets are the simplest form of panel data. I alleviate the selection bias due to unobserved characteristics by comparing how much an individual does when he/she uses contacts with how he/she does when he/she does not use contacts. The within-subject comparison rules out any time-fixed individual unobservables on which the selection of job-seeking method may arise. In this study, I thus allow the unobserved individual factors that affect the job outcomes to also be correlated with whether the individual uses contacts or not.

Assuming that individual unobservables influencing wages (α_i) are constant over time from *t* to *t-1*, one can differentiate the data across two time-points/waves to rule out the time-fixed unobservables. I first write two separate cross-sectional wage equations for two years, namely year *t* and year *(t–1)* as:

$$W_{it} = \beta_{0t} + \beta_1 C_{it} + \beta_2 X_{it} + \alpha_{it} + \epsilon_{it} \quad (2)$$

$$W_{i(t-1)} = \beta_{0(t-1)} + \beta_1 C_{i(t-1)} + \beta_2 X_{i(t-1)} + \alpha_{i(t-1)} + \epsilon_{i(t-1)} \quad (3)$$

Then I subtract Eq. (3) from Eq. (2) and obtain:

$$W_{it} - W_{i(t-1)} = \left(\beta_{0t} - \beta_{0(t-1)}\right) + \beta_1\left(C_{it} - C_{i(t-1)}\right) + \beta_2\left(X_{it} - X_{i(t-1)}\right) + \left(\alpha_{it} - \alpha_{i(t-1)}\right) + \left(\epsilon_{it} - \epsilon_{i(t-1)}\right) \quad (4)$$

Since I assume that an individual unobservable α_i is time-invariant, that is, $\alpha_{it}=\alpha_{i(t-1)}$, Eq. (4) ends up as:

$$\Delta W_i = \alpha_0 + \beta_1 \Delta C_i + \beta_2 \Delta X_i + \Delta \epsilon_i \quad (5)$$

where "Δ" denotes the change from t to (t–1). Notice that α_i disappears from the equation as α_{it} is equal to $\alpha_{i(t-1)}$. Eq. (5) is therefore the first-difference model, predicting changes in wages as a function of changes in the independent variables. Since the first-difference models estimate longitudinal models of multiple job searches for the same respondent (in this study two job searches), I only use individuals who changed their jobs at least once during the two time periods.

4. RESULTS FROM THE GDR DATA

A significant advantage of panel data is that the questions of interest (in this study the questions about the use of contacts to find work and job outcomes) were asked repeatedly at each wave of the survey. The GDR data collected by Völker and Flap (1999) provide us with an opportunity to exploit the richness of data in a panel analysis. Notice that, although Völker and Flap (1999) collected two-period panel data (first job and job in 1989), they actually fitted two OLS

models respectively for the two time periods and report that there is no evidence supporting the direct effects of using contacts on the respondent's occupational prestige. In this section, I first replicate their study and then proceed to extend their study by fitting a first-difference model.

The GDR data were collected in 1992, 1993 and 1994 in two cities, Leipzig and Dresden. In this study, I focus on the data referring to 1992 and 1993, as the two surveys asked retrospective questions about job and social networks in the socialist period (i.e. first job and job in 1989). For a detailed description of the design of the study and the data collection, one can refer to Völker (1995). Using the unique data set, Völker and Flap (1999) regress the relevant job outcomes (i.e. the occupational prestige of the first job, and the occupational prestige of the job in 1989) on the uses of contacts and a set of individual controls (i.e. father's education, father's occupational prestige, the education level of the respondents and their sex). To replicate their study, I use the same set of explanatory variables as Völker and Flap (1999)[62]. Table 2 above presents the descriptive statistics of the GDR data. Notice that, as this study uses age to proxy the experiences as the control in one of the first-difference models, the sample size (483) is slightly different from that of Völker and Flap's (1999) original, which is 488. Also, since the first-difference models only use workers who had changed jobs (i.e. jobs held in 1989 were different from their first jobs), I distinguish the job-changers and

[62] Education is included in the model as a continual variable rather than as a dummy variable. This is so as to be consistent with the model specification in Völker and Flap (1999).

all cases in Table 2. As can be seen, 51 out of the 483 workers stayed in the same work unit as their first job, while 432 workers changed positions.

Table 2: Selected Descriptive Statistics of the GDR Data (N = 483)

	All Cases N = 483		Job-changers N = 432	
Individual attributes	***Mean***	***S.D.***	***Mean***	***S.D.***
Age	41.38	8.02	41.38	8.01
Occupational prestige (SEI first job)	47.28	12.63	47.18	12.42
Occupational prestige (SEI job in 1989)	47.82	13.21	47.80	13.07
Occupational prestige of father (SEI)	44.91	12.24	44.75	12.16
	Freq.	**Percentage**	**Freq.**	**Percentage**
Use of contacts (first job)	116	24.02%	101	23.38%
Use of contacts (job in 1989)	195	37.04%	176	40.74%
Female	268	55.49%	239	55.32%
Education 1 (father)	66	13.66%	64	14.81%
Education 2 (father)	106	21.95%	86	19.91%
Education 3 (father)	89	18.43%	81	18.75%
Education 4 (father)	132	27.33%	125	28.94%
Education 5 (father)	47	9.73%	39	9.03%
Education 6 (father)	43	8.90%	37	8.56%
Education 1	16	3.31%	15	3.47%
Education 2	69	14.29%	62	14.35%
Education 3	152	31.47%	136	31.48%
Education 4	26	5.38%	23	5.32%
Education 5	93	19.25%	81	18.75%
Education 6	127	26.29%	115	26.62%

Notes: (1) Occupational prestige is measured by the Treiman occupational prestige scale for the German version developed with the National Bureau for Statistics.
(2) Education 1 = finished school before 8th grade; Education 2 = high school until 8th grade; Education 3 = high school until 10th grade; Education 4 = high school (university qualification); Education 5 = technical college; Education 6 = finished university.
(3) SEI = Social Economic Index

Table 3 below presents the empirical results obtained from both the OLS and the first-difference models[63]. In particular, Model 1 and Model 2 in Table 3 replicate the results of Völker and Flap (1999: 24, Table 5) except for two things: 1) I use a dummy coded as 1 or 0 to denote the use of contacts, while they include two dummy variables in the model; and 2) the error terms in my model are robust to heteroscedasticy. In the two OLS models referring to the first job (Model 1) and the job held in 1989 (Model 2), both of the estimated effects of using contacts on occupational prestige are almost identical to those in Völker and Flap (1999: 24). The results obviously indicate that using contacts to find jobs does not influence the status attainment. As for the other controls, the estimated coefficients are also very similar to those used by Völker and Flap (1999). Notice that, since there are 51 workers that did not change jobs, their first jobs are the same as their 1989 jobs. This problem is dealt with in Model 3.

Since the first-difference model is to be estimated on job-changers, Model 3 uses job-changers for comparison in order to conduct the OLS analysis. The age is also included in Model 3 as it proxies experiences, which is time-variant and expected to influence job outcomes. As can be seen from Column 3, estimates from Model 3 indicate that using contacts does not have significant effects on the occupational prestige of the job-changers, given the 0.05 alpha level. Obviously

[63] Table 3 presents the estimated coefficients and the standard error.

this is also consistent with the estimated effects obtained from Model 2.

Table 3: Models of Occupational Prestige from the GDR Data

	OLS			First-difference
Independent Variables	**Model 1** **All cases** **First job**	**Model 2** **All cases** **Job 1989**	**Model 3** **Job-changers** **Job 1989**	**Model 4** **Job-changers**
Use of contacts	–1.01	1.75*	1.88*	.41
	(1.132)	(0.991)	(1.073)	(1.053)
Female	4.39***	–.16	-.13	–
	(0.975)	(0.98)	(0.903)	–
Education	3.96***	2.31***	2.55***	–
	(0.330)	(0.411)	(0.427)	–
Father's education	.611	-.038	-.051	–
	(0.395)	(0.396)	(0.422)	–
Prestige of father's job	.0578	.050	.091*	–
	(0.048)	(0.047)	(0.051)	–
Prestige of first job	–	.44***	.342***	–
		(0.056)	(0.060)	–
Age	–	–	-.431	-.519
	–	–	(0.811)	(0.971)
Age-squared			.0069	.0074
			(0.009)	(0.012)
Constant	24.59***	15.33***	22.24	1.29
	(1.971)	(2.515)	(16.79)	(5.09)
R-squared	0.31	0.39	0.35	0.007
Number of observations	483	483	432	864
Number of cases	–	–		432

Note: i) Standard errors are in parentheses and robust to heteroscedasticy.
ii) * p < = 0.1; ** p <= 0.05; *** = p < 0.01
iii) Model 1 and Model 2 include all cases; Model 3 and Model 4 include only those respondents who changed jobs.

Notice that, although the GDR data do not contain information regarding the precise age that the respondents started their first jobs, I assume that the age

when starting the first job is 18 for those without university degrees and 22 for those with university degrees[64]. As can be seen from the last column in Table 3, the estimated coefficient of the effects of using contacts decreased from 1.88 to 0.41 and is not significant anymore, even given a 0.1 significance level. Notice that all the effects of time-fixed observed factors such as education and gender are ruled out by differencing. In general, the first-difference estimates demonstrate that the use of contacts does not matter for the occupational prestige of the secured job. Compared to the OLS estimates, the first-difference estimate is more meaningful because it has ruled out the time-fixed individual heterogeneity. If we believe that the omitted variable (e.g. individual ability, social class or occupational type of the respondents) on which selection may arise is consistent from his/her first job to that held in 1989, we can conclude that using contacts does not have any causal effect on the occupational prestige of workers in the GDR.

However, it is possible that the temporal interval between the first job and the job held in 1989 may be very long. In this case, it is not reasonable to assume that individual heterogeneity can be time-fixed. To explore this, I also use those who are younger than 35 years and younger than 30 to fit the model. Although not reported in Table 3 for reasons of space, the results are basically the same: we see no significant effects from using contacts. Nevertheless, in the following

[64] Although not reported here, I also tried excluding age and age-squared as controls and got very similar estimates.

section analysing the Chinese data, the interval between the two time periods is just three years. As a result, one can more confidently assume that the omitted individual factors are time-invariant during that period.

5. RESULTS FROM THE CHINESE DATA

Although there has been a steady accumulation of literature regarding job searches and social networks in the Chinese labour market, little attention has been paid to the direct role of using contacts. This section looks into whether the GDR pattern applies to China, which is of course also governed by state socialism. If China's case does not support the efficacy of using contacts, we would be able to more confidently conclude that the use of contacts does not by itself matter under state socialism. That said, the use of contacts to find a job is very common in China. Based on a survey interviewing urban residents in Tianjin City, Bian (1997) reports that roughly 45 per cent of Chinese urban residents found their jobs through informal channels. Using samples from 18 Chinese cities, Lin (2003) reports that around 15 per cent of their inhabitants use personal contacts, although he stresses that the percentages are substantially underreported, as the use of contacts is a "behaviour much more often done than spoken of in public" given that the labour markets are to a large extent supervised and dictated by governments (Lin 2003: 152).

In order to carry out the present analysis, I exploit the richness of the data from the 2002 CHIPS. This survey was jointly conducted by the Urban Survey Group of the National Bureau of Statistics of China (NBSC) and the Institute of Economics of the Chinese Academy of Social Science (CASS). The data used in this study are from a sample of 20,632 urban residents from 22 provinces[65]. The questionnaires were collected in February 2003 and gathered detailed information on the demographics, recent employment history and wages of the current and previous jobs of urban respondents. This offers the opportunity to do a panel analysis to probe the wage effect of using contacts. As for the job-seeking method, the questionnaire used in CHIPS 2002 asked respondents to provide detailed information regarding the respondents' current and previous jobs. With regards to job-search method, the exact phrasing of the question reads: "How did you get your current job?" and "How did you get your previous job". Ten possible answers are offered: 1) Assigned by government; 2) Inherited it; 3) Through an open examination; 4) Through an employment agency; 5) Through a newspaper; 6) Through a referral by someone you or your family knows; 7) Found it on your own; 8) Started your own business; 9) Not working currently; 10) Other. Only one answer can be chosen. If the respondent chooses the sixth answer, he/she is classified into the group of using contacts.

[65] The sample does not include rural-to-urban migrants.

To conduct a comparative analysis between the GDR and Chinese cases, it would be ideal focus on the same dependant variables. This study, however, looks into the role of using contacts on wages rather than on job status. The reason is straightforward. Notice that the definitive feature of rankings of occupational prestige is the stability stemming from variations in structural positions, social institutions, and cultural contexts (Zhou 2005). But China has been undergoing dramatic social-economic changes from 1980s. As a result, the studies of Chinese occupational prestige have encountered difficulties due to the historical changes during the decades (Bian 1996). For example, Lin and Xie (1988) found that although rating of occupational prestige by urban residents in China is consistent with those found in other countries, specific differences are found suggesting that recent historical experience influences evaluations on job status. Bian (1996) reports the university of occupational prestige, but he finds that the surrounding political atmosphere and the hidden benefits of jobs under planned market economy affect the rating of job status among Chinese. In a later study based on Beijing survey, Li (2000) even reports conflicting occupational prestige evaluation between Chinese and people in other countries. More importantly, although so far sociologists have developed several scales to rate job status in China(see e.g. Li 2005; Xu 2000), few of them are well established and recognized.

As previously mentioned, it is not reasonable to assume that individual

heterogeneity can be time-fixed for a long time. As a result, subtracting the wage equation of the first job from the wage equation of the current job may not rule out individual heterogeneity. However, one can comfortably assume that the omitted individual factors are time-invariant during the relatively short period of three years. In the 2002 CHIPS, the respondents were asked the question "Have you changed jobs in the last three years?" [66] This allows me to compare the different results between using contacts and not using contacts for the same respondent within a three-year period. Notice that, in estimating the first-difference model, those who had not changed their job between 2000 and 2002 will not be used, because their "previous job" is the same as their "current job" within the relevant period that the 2002 CHIPS asks about.[67] In addition, respondents younger than 16 and older than 60 have been dropped from the sample. Considering the potential outliers in terms of wages, I also dropped 236 respondents who reported zero wages.[68] Given the missing data of other variables of interest, the final size of the useful working sample is 9,328. Among them, there are 750 job-changers who had changed job during the three years in question.

66 This refers to the period 2000 to 2002.

67 Note that, from now on, "previous job" refers to the job before the current job after 1999, although many non-changers (1999–2001) did have previous jobs before 1999.

68 Zero wage led by wages arrears are rare among urban Chinese who are employed, although it is not unusual among rural-to-urban migrants. In robustness checking, I also dropped observations in the top 1% and bottom 1 % from the sample. The results are very stable.

Table 4: Selected Descriptive Statistics of the CHIPS Data

	All Cases N=9328		Job-Changers (in 3 years) N=750	
Number				
	Mean	***S.D.***	***Mean***	***S.D.***
Age	41.34	8.27	38.6	8.15
Years of schooling	11.39	2.99	11.1	2.94
Annual wage of previous job*	--	--	7506	6062
Annual wage of current job	12493	8971	10271	11292
Experience for previous job	--	--	21.23	9.53
Experience for current job	23.95	9.38	21.52	9.42
	Freq.	**Percentage**	**Freq.**	**Percentage**
Female	4060	43.52%	364	48.53%
Using contacts (previous job)	--	--	121	16.13%
Using contacts (current job)	754	8.08%	202	26.93%
Work unit sector (current job)				
Governments/institutes	2945	31.57%	159	21.20%
State-owned enterprises	3210	34.41%	107	14.27%
Collectives	602	6.45%	34	4.53%
Joint ventures	173	1.85%	17	2.27%
Private/self-employed	2398	25.71%	433	57.73%
Work unit sector (previous job)				
Governments/institutes	--	--	96	12.80%
State-owned enterprises	--	--	217	42.27%
Collectives	--	--	98	13.07%
Joint ventures	--	--	12	1.60%
Private/self-employed	--	--	227	30.27%

** The "previous job" refers to the prior job within the previous three years.

In Table 4, the first column describes the full sample while the second column describes the characteristics of the respondents who changed jobs during the previous three years. As can be seen, when respondents were asked to report the

methods of job search they used for their current job, around 27 per cent of job-changers claimed that they used contacts to find jobs, while only 8 per cent of the full sample reported using contacts. This implies that recent job-changers tend to use informal networks. This might be due to the fact that, from the late 1990s onwards, a "real" labour market started to emerge in China[69]. From then on, people could be expected to be more likely to use informal networks to find or change jobs. As for demographic characteristics, recent job-changers in general are relatively younger, are less educated, earn less, and have less experience. Among job-changers, we can see that, during the three years 2000 to 2002, many workers in state-owned enterprises turned to self-employment.

In the data analysis of the GDR study, I showed that the OLS model on the full sample, the OLS model on job-changers, and the first-difference model on job-changers yielded different results. Moreover, based on the first-difference estimates, we saw that the use of contacts does not have significant effects on occupational prestige. To see whether this pattern can also be found in China, I proceed to estimate three wage models on the data of 2002 CHIPS. To reduce positive skew in the data of wage, I use logarithmic transformation. The results are presented in Table 5.

[69] Another possible explanation could be that the question on job search methods was only asked of those who actively searched for work.

Table 5: Models of Wage of Current Job Using CHIPS Data

	OLS	OLS	First-difference
	Model 1	Model 2	Model 3
Independent variables	All cases	Job-changers	Job-changers
Use of contacts	–.141***	–.136***	–.080
	(.023)	(.051)	(.052)
Female	–.181***	–.057	–
	(.011)	(.045)	–
Education	.064***	.042***	–
	(.002)	(.010)	–
Experience	.026***	.022	–.028
	(.003)	(.010)	(.072)
Experience-squared	–.0003***	–.00004	.0002
	(.000)	(.0002)	(.002)
Work unit sector			
State-owned enterprises	–.185***	–.041*	.060
	(.013)	(.021)	(.056)
Collective	–.457***	–.195**	.195**
	(.024)	(.092)	(.087)
Joint venture	–.029	–.077	–.006
	(.047)	(.27)	(.183)
Private/self-employed	–.321***	–.105*	.185**
	(.024)	(.056)	(.002)
Wage of previous job	--	.48***	--
	--	(0.045)	--
Province dummies	YES***	YES***	–
Constant	8.954***	4.627***	.255***
	(.056)	(.487)	(.028)
R-squared	0.32	0.38	0.03
Number of observations	9328	750	1500
Number of cases	–	–	750

Note: i) Standard errors are in parentheses and robust to heteroscedasticy.
ii) * p < = 0.1; ** p <= 0.05; *** = p < 0.01
iii) Model 1 includes all cases; Models 2 and 3 include job-changers only.

Model 1 estimates the effects of using contacts on the wages of the current job

among all of the 9,328 respondents. The estimated coefficient of using contacts is –0.14 with a standard error of 0.023, indicating that using contacts in fact decreases the wage. Specifically, all else being equal, compared to the wages of those who used contacts to find the current job, the wages of those who had not used contacts are 14 per cent higher. Although seemingly counter-intuitive, this is consistent with some previous findings (e.g. Smith 2000; Bentolila, Michelacci, and Suarez 2010; Delattre and Sabatie 2007; Ponzo and Scoppa 2009). As for other controls, as expected, urban females have lower wages in comparison with males. The return to education is also significant and one more year of schooling leads to 6.4 per cent higher log wages. Experience also matters in terms of wages, and the relationship between experience and wage is non-linear, as experience-squared also significantly predicts wages. When it comes to work unit, compared to party cadres and governmental officials, those who work within state enterprises, collectives or private enterprises have lower wages. However, this pattern does not apply to employees in joint ventures or foreign-owned enterprises. Finally, the significant effects of provincial dummies reflect the regional differences in terms of economic development in contemporary China.

Before proceeding to fit the first-difference model, which explores the variation in job-search methods for the same person between his/her current and previous job, I first of all estimate an OLS model, namely Model 2, for recent job-changers. The sample size thus shrinks dramatically from 9,328 to 750,

because only a small fraction of people changed their position in the previous three years. I also include the variable of wages of previous jobs to control for some unobserved factors. The estimated effects of using contacts, however, seem to be unaffected by the lost sample size. As can be seen in the second column in Table 5 above, the coefficient of the contacts variable in Model 2 is –0.18, which is also significant given a 0.05 alpha level. This is quite close to the OLS estimate of the effects of using contacts on the full sample in Model 1. This pattern also applies to most of the other control variables. The changes lie in the fact that the effects of gender and experience disappeared.

The third column presents the results obtained from the first-difference model. I found that, when time-fixed individual heterogeneity is taken into account, the effects of using contacts substantially increased by more than 50% (from –.18 to –.079) and are not significant any more at the 0.05 alpha level. This implies that the significant negative role of using contacts suggested by Model 1 and Model 2 may be spurious and actually reflects the roles of some time-fixed individual characteristics. For example, if people of lower earning ability tend to use referrals to find a job, or lower-wage occupations are more likely to be filled by informal networks, then the role of using contacts will be biased downward if we do not take into account the individual ability or occupational types which were omitted from the OLS model. When the first-difference model ruled out the time-invariant heterogeneity, the estimated role of using contacts increased

because the effect of using contacts has not been mistakenly offset by the confounder effects. This pattern has actually also been found in Mouw (2002). In his study of the NLSY panel data, the estimated effects obtained from a fixed-effects model are positive, while the OLS estimates are on the contrary negative.

6. FURTHER DISCUSSION

Taking into account the time-fixed unobserved characteristics that may cause selection, results from the first-difference model suggest that using contacts does not pay off. Furthermore, comparing the GDR and Chinese cases, I found that selection influences the estimation in different directions under different social contexts. Specifically, results based on the GDR data suggest that the status effect of using contacts obtained from the OLS model is upwardly biased, while results based on CHIPS data strongly imply that the OLS estimate of wage effect is downwardly biased. This means that there could be different social mechanisms through which job-seeking methods are selected, even under ostensibly similar political systems. To further investigate this issue, I propose two different possible forms of selections during the process of choosing job-search methods to interpret the empirical findings in the GDR and China.

The first one is, as I call it, positive selection. That is, using contacts is correlated

with wages and the omitted/unobserved variables not only at the same time but also in the same direction. A typical positive selection is caused if more able people (in terms of status attainment or earning, and unobserved) tend to use contacts to find jobs. In this case, the effects of using contacts are biased upward when naively fitting an OLS model. Notice that positive selection can also arise if high-wage jobs are more likely to be filled by those using contacts and industry variables are omitted from the model. However, since most previous empirical studies report that it is poor jobs that are more likely to be filled by those using informal networks (see e.g. Holzer 1987; Marsden 2001; Corcoran et al. 1980), it is not discussed here.

As we can see, the GDR case demonstrates positive selection very well. Obviously, it is natural to believe that this pattern is more likely to occur if using contacts is not a taboo. Within such a social context, those of higher ability would tend to use contacts. For the GDR in 1989, the job-assignment system meant that everybody got a job, and that there was no genuine labour market in which formal channels such as agencies could be used to find jobs. Thus, more able workers are more likely to turn to social ties to get better jobs. To check for this, I proceed to examine whether the German respondents with observed characteristics that signal higher ability (e.g. male, more experienced, more educated) are more likely to use contacts. For consistency of analysis, I only examine job-changers.

As can be seen in Table 6 below, the average level of education is higher among contact users (4.08) in the GDR than among non-contact users (3.96). Similarly, the average age of contact users (41.65) is also higher than that of non-contact users (41.18). In addition, a probit model predicting a dichotomous variable indicating the use of contacts was estimated, although in order to save space the results are not reported here. Although this equation explains limited variations in the use of contacts, and the explanatory variables (i.e. age, age-squared, gender, education, father's education and father's job) have no significant effects on the use of contacts, the coefficients are positive and the results in general suggest that contacts tend to be used by those who are in better socio-economic positions.

Table 6: Comparing Contact Users and Non-Contact Users (GDR Data)

	Contact users		Non-contact users	
Number	176		256	
Ability-related characteristics	Mean	S.D.	Mean	S.D.
Education	4.08	1.60	3.96	1.54
Age	41.65	8.08	41.18	7.96

The other type of selection is called negative selection: using contacts is correlated with wages and the omitted/unobserved variables at the same time but in different directions. For example, less able people tend to use contacts to

find jobs, or jobs filled using contacts are concentrated in low-wage industries. In the presence of negative selection, the first-difference estimates would be larger than the OLS estimates because the "genuine" effects of using contacts would have been mistakenly offset by the effects of unobserved individual ability or omitted industries. Three potential sources of negative selection are discussed below to explore which ones are likely to be the Chinese case.

Firstly, negative selection can be driven by unobserved individual ability. It is easy to see that this mechanism works when the use of contacts is seen as "shameful", so that people of higher ability are consequently more likely to use formal ties as long as formal channels are available. For China in early 2000s, for example, the job-assignment system had been ended and labour markets were in the making. Although using contacts in job-seeking was becoming less unacceptable, it remained a mode of behaviour that, although prevalent, was not spoken of in public (Lin 2003). In addition, negative selection driven by ability could also be the result when the using of contacts engenders high reciprocal obligations. In this case, as rational actors evaluating the cost and the expected return of actions, people with higher ability will first try formal job-search channels such as direct applications, entry examinations or job agencies before they turn to informal networks. This may also be the case in China in early 2000s, as mobilising one's social resources or using social contacts is accompanied by moral and economic obligations (Hwang 1987), and the emerging labour

markets made formal channels for jobs available..

As with the GDR analysis, I also examined the 2002 CHIPS data to see whether Chinese respondents with observed characteristics that signal less ability, poor employment opportunities or low-wage industries (e.g. women or the less educated) are more likely to use contacts. As can be seen in Table 7, I did find a systematic pattern in the selection of job-searching methods. For example, the average years of schooling among contact users is roughly 10, while the average education among non-users is 1.5 years more. As for age, contact users are on average two years younger than non-users. I also estimated a probit model predicting the use of contacts. Although not reported here, the results suggest that using contacts is significantly and negatively correlated with education.

Table 7: Comparing Contact Users and Non-Users (Chinese Data)

	Contact users		Non-users	
Number	202		548	
Ability-related characteristics	Mean	S.D.	Mean	S.D.
Years of schooling	10.52	2.56	11.31	3.05
Age	38.9	8.43	38.5	8.05

Secondly, negative selection can be caused by unobserved job offer opportunities. Montgomery (1992) was among the first to propose this possibility in modelling

network effects. The major hypothesis is that people will use contacts only when the job offer is better than those produced by formal channels (Montgomery 1992). If the level of job-offer opportunity were omitted from the model, because the opportunity is positively correlated with wage and negatively correlated with the using of contacts, it would lead to an underestimated effect using the standard OLS model. In this case, as Mouw (2002) puts it, a negative relationship between contacts and wages in cross-sectional data may simply reflect a lower level of opportunity in general, rather than a particular negative effect arising from the use of job contacts. Since the data from 2002 CHIPS has not provided us with any information of labour market conditions of respondents, it is likely that the estimation bias can at least be partly ascribed to the unobserved job offer opportunities.

Thirdly, negative selection is likely to be led by omitted industries. If jobs filled using contacts are concentrated in low-wage industries are not included in the model, the standard estimates would be downwardly biased as well. This pattern has been found in numerous empirical studies. For example, researchers have found that low-wage jobs (Holzer 1987) and non-professional jobs (Marsden 2001) tend to be filled via employee referrals and that blue-collar workers are more likely than white-collar workers to use contacts (Corcoran et al. 1980; Marsden 2001). However, since dummy variables of sector of work unit of respondents are controlled for in the wage model on Chinese data, negative selection found here is less likely to be driven by omitted industries. Notice that

the above three potential sources of negative selection may work at the same time. For instance, in a cross-border study in Europe, Pellizzari (2004) found that, in most countries, occupation and education (which partly reflect ability) are both negatively correlated with the use of contacts to find jobs. He goes so far as to call it a cross-country regularity.

No matter how the selection of job-seeking strategies biases the OLS estimates in different directions, it is still of interest to ask why using contacts does not matter. There are several potential explanations. For example, if using contacts is a common behaviour in the process of job-seeking, but only some people reported using contacts, then the direct role of using contacts would be marginal. Although we found in previous studies that the percentage of respondents using contacts ranges from 15% to 50%, this could be underreported because using contacts is seen as an "unwritten strategy" in many social contexts. Another explanation may be that using contacts alone does not matter in terms of job outcomes, but, rather, it is the characteristics of the contacts, or the mobilised social resources, that influence job outcomes since it is the job-related information which affects job outcomes rather than the simple use of contacts by itself (see Mouw 2003 for further discussion on this). This proposition will be examined in Chapter 6.

7. CONCLUSION

Informal networks are widely used by people to find jobs. However, so far it has not become clear whether using contacts leads to better job outcomes than not using contacts. People may select their particular job-seeking method according to certain unobserved individual factors or labour market conditions. This suggests that the standard OLS estimates are threatened by a potential endogeneity problem in estimating the direct role of using contacts. To address the selection bias, one needs more informative data and more advanced models than naive OLS models offer. However, due to data limitations, sociological analysis of the effect of using contacts has not paid sufficient attention to this issue. In addition, most previous studies focus on labour markets under market economies. Little light has been shed on the job-seeking process under state socialism. Consequently, this study contributes to the literature by exploiting the first-difference model to get around the selection bias and extending this line of research to countries under state socialism.

Under the working assumption that the unobservable characteristics are invariant over time, I obtained consistent estimates from the first-difference models for both the GDR and Chinese data. In particular, I first replicated and extended Völker and Flap's research (1999) on urban residents in the GDR. Then I used the same model on the data from the 2002 CHIPS. Both the GDR and

Chinese results suggest that there is no evidence supporting the notion that using contacts matters in the two countries under study. However, according to the GDR study, the effect of using contacts on occupational prestige is upwardly biased if one does not take into account potential selection. In contrast, the Chinese study demonstrates that the OLS estimate of the effect of using contacts on wages is downwardly biased. Based on this study, the picture is that the GDR sees positive selection while China sees negative selection. Potential sources and mechanisms of these types of selection behaviour are discussed.

On balance, to estimate the payoff of using contacts compared to not using contacts, we need an understanding of how people actually select their job-seeking strategies. This class of selection is deeply rooted in the specific social contexts that we are looking at and calls for more attention in future studies. I would not argue that using contacts does not matter based on this study. Rather, I would say that the empirical findings here confirm the argument of Ioannides and Loury (2004) that the effects of informal networks on individual outcomes may differ according to the types of contacts, referrals, work environment, institutional characteristics, and so on. In particular, we can conclude that selection derived from unobservables may be different between countries, even those under the same political system.

Chapter 6

Do Contacts Matter under State Socialism? A Replication and Extension

1. INTRODUCTION

Sociologists have long been interested in the impact that the resources embedded in social structure have on individual outcomes in the labour market. The theoretical intuition is straightforward: social ties or networks provide access to job-related information and influences. As a result, those who have used better resources (e.g. high occupational prestige of contacts, wide-ranging networks, or well-educated friends) will obtain better socioeconomic outcomes (e.g. status, wages, or job satisfaction). Lin, Ensel and Vaughn (1981) were among the first to take up this theme by empirically studying the ways in which embedded resources, or, social capital, influence job matching. Since Granovetter's (1985) ground-breaking contribution stressing the role of weak ties, research focusing on network effects and social capital has increased. Among the early studies, a research tradition has been established to identify how the social capital used by the focal individual to find jobs affects the individual's job outcomes (Bridges and Villemez 1986; Marsden and Hurlbert

1988; Völker and Flap 1999; see Lin 2008 for a review). In general, evidence from these analyses is in favour of the significant role of the used social capital.

However, that being said, if not all, most of the early studies have inadequately addressed the endogeneity problem plaguing casual inference. The major sources of the endogeneity are twofold: first, the use of contacts to find a job may be selective (as the preceding chapter reveals) and, if so, the samples that sociologists look at are unrepresentative. Second, there might be unobserved factors (e.g. earning ability, personality, etc.) influencing both the social capital used (e.g. contacts' prestige) and the individual outcome of interest. In this case, the estimation of the standard OLS model will be biased and the speculated causality between social capital and individual outcomes may be spurious. For example, Mouw (2003) replicates the study by Marsden and Hurlbert (1988) and argues that it is actually the social homophily – the similarity between the occupations of the job seeker and the contact – rather than the prestige of the contact that matters.

Paying close attention to the endogeneity problem, this study presents an extension of this line of studies of Lin, Ensel, and Vaughn (1981), Marsden and Hurlbert (1988), Völker and Flap 1999, and Mouw (2003). To deal with the potential sample selection problem, I exploit the Heckman two-stage model to allow for an unrepresentative sample. To correct for other omitted variable

biases, I follow the same respondents over time to see whether any significant difference develops between the time when they used contacts for the previous jobs and used contacts for the current jobs in the surveying year. The data collected and used by Völker and Flap (1999) is re-analysed to identify the effects of social capital in the GDR. The data from 2002 CHIPS is also examined to obtain more evidence under state socialism. The replication on the GDR data confirms the major findings of previous studies: the contacts' occupational prestige matters for the workers' occupational prestige. The evidence from China also shows that there is a causal relationship between the contacts' membership of the Chinese Communist Party (CCP) – political status as a specific type of social capital – and respondents' wages.

The rest of the paper is structured as follows: Section 2 provides relevant background on the endogeneity problem which affects the identification of the causal effect of social capital in this line of research. Section 3 addresses some of the different models and identification strategies that are to be used to correct for potential endogeneity biases. Section 4 analyses the GDR data and Section 5 looks in detail at the Chinese data. Finally, Section 6 presents some concluding remarks.

2. BACKGROUND

There has been a large body of empirical studies on the role of social capital. Central to this line of research enterprises is Lin's (1999) social resource theory. It argues that the better or more social capital is used, the better job outcomes will be. It is very appealing as we know that social resources are unevenly distributed in the social ladder. As a result, mobilising the resources possessed by others pays off. In these studies (e.g. Lin, Vaugh and Ensel 1981; Lin, Ensel and Vaugh 1981; Marsden and Hurlbert 1988; DeGraaf and Flap 1988; Requena 1991; Wegener 1991; Bian and Ang 1997; Bian 1997; Völker and Flap 1999:42; Smith 2000; Mouw 2003; Lin 2003; Lin and Ao 2008), the social capital used is often operationalised by the occupational prestige of the job contacts. As for the job outcomes, status and wages are typically used. Lin (2008) presents a cogent review of these studies. According to him, a consensus has been reached that high status job helpers are beneficial to job outcomes. Almost all of the above studies, however, are based on a cross-sectional analysis and the OLS model. This, therefore, has left the omitted variable and selection problem open[70] and leads to two sources of potential endogeneity bias.

The first source of endogeneity is the conventional omitted variable bias which is

70 Since most of the related survey questionnaires explicitly asked the respondents to provide information on the specific job helper (through questions such as: "what was the occupation of the job helper when they helped you to get your current job?"), the time order is clear. This means that there is no mutual causality problem. That is, the simultaneity problem which is often another source of endogeneity will not arise.

very similar to the "ability bias" in econometric studies on the return to education (see Card 1998 for a review). For instance, the measure of individual ability is often crude or unavailable in the data sets that social scientists work with. As a result, it is often omitted from the standard OLS model. Since ability is correlated with both educational attainment and earnings, the estimated educational effect would be upwardly biased, because it also reflects the role of individual ability. In the case of a social capital study, if the individual heterogeneity influences both making friends or finding contacts and getting jobs, the role of social capital will be biased as well. For example, if people of high ability tend to make friends with, or turn to those who are similar to them for help, the unmeasured individual ability would be associated with the measure of social capital as the person's contacts or friends are also of high ability and thereby high occupational prestige. If so, the explanatory variable, the social capital (e.g. the occupational prestige of contacts) is correlated with the error term in the model, leading to biased estimates.

Aside from the omitted variable problem, sample selection is the other leading source of endogeneity. For example, individuals of higher social status are less likely to use contacts (Marsden 2001). As a result, when one regresses the job status on the social capital, using a "truncated sample" this leads to a biased estimation. Another example of this is that previous studies demonstrate that richer human capital predicts a higher probability of using contacts to find work

(see e.g. Bentolia et al. 2010; Lin and Ao 2008). If human capital is not controlled for, it actually enters into the error terms of both the substantive model and the selection model, thereby giving rise to biased estimation. As Marsden and Hurlbert (1988) stress in estimating the causal effect of social capital in the labour market, there are substantive reasons to think that unmeasured factors predisposing people to use contacts may be correlated with error terms in the job outcome equations.

Mouw (2003) finds that when the similarity of occupation between the job-seeker and the job-helper is controlled for, the significant effect of used social capital disappears. He then concludes that: "the evidence in favor of the social resources perspective is largely an artifact of the incidence of same-occupation information flows between contacts and job seekers" (Mouw 2003:883). Although Mouw (2003) operationalises the social homophily factor (i.e. the similarity of occupation) in a problematic way, his argument that the used social resources might be endogenous is perceptive. However, few sociological studies have furthered this issue to systematically deal with the potential endogeneity problem.

To be fair, it should be noted that the early studies are not totally unaware of the potential endogeneity problem and, in general, used two methods to deal with it. The first is to include the time-lagged dependent into the model. In particular,

since the “prior job” or the “first job” had been widely used as controls, previous studies had in fact employed the lagged term of dependant as a proxy for unobserved factors. This method is reasonable but far from perfect. This is because to proxy an unobserved individual factor using job outcomes from a previous period (immediate prior job or first job) introduces a number of additional unobservable factors into the model. The other method is using the Heckit approach to deal with sample selection(see e.g. Marsden and Hurlbert 1988; Wegener 1991; Lin 2003). However, as reviewed in previous chapters, these studies have not sufficiently imposed exclusion restrictions on the selection model. As a result, their empirical findings are still open to the sample-selection bias.

On balance then, so far sociological analysis of the effects of used social capital is far from satisfactory. Firstly, limited data has led to a lack of model identification strategies. Although early studies report a positive association between social capital and individual outcomes, this relationship cannot be seen as causal due to the potential endogeneity biases. Secondly, although recent studies such as Mouw (2003) have been aware of the endogeneity problem, to the best knowledge of the author the empirical studies exploiting richer data and advanced models are very rare. Thirdly, the existing literature on the effects of used social capital is mainly based on free market economies. Relevant research under state socialism is scarce and has not paid enough attention to the

endogeneity problem (e.g. Völker and Flap 1999; Bian and Ang 1997; Lin 2003). This chapter, consequently, aims to fill this gap and to shed more light on how to empirically identify the causal role of used social capital under state socialism.

3. MODELS

The basic specification of the model includes the major explanatory variable of interest, the resource possessed by the contact (job helper), the observed individual characteristics as controls (e.g. age, age-squared, sex, educational attainment, etc.), and unobserved individual heterogeneity. Taking the wage as the dependent variable, for example, to estimate the effect of used social capital on wages, I start with a standard OLS model which can be expressed by:

$$W_i = \beta_0 + \beta_1 S_i + \beta_2 X_i + \alpha_i + \epsilon_i \quad (1)$$

Where W_i is the wage of the individual *i*. S_i denotes the level of the used social capital, X_i denotes a vector of individual factors, α_i is the individual unobservable influencing wages and ϵ_i is the random error. In Eq. (1), β_1 is the estimation of the social capital effect. This method, however, is confronted by a major difficulty. Notice that one of the most important premises for obtaining an unbiased OLS estimate of β_1 is that Cov $[S_i . \alpha_i] = 0$. That is, the variable representing resources possessed by a given contact and used by the focal individual must be exogenous.

However, this assumption is very unlikely to hold in the context of this study because of the potential presence of sample selection and other conventional endogeneity problems. As a result, alternative models and richer data are needed.

The Heckit model that is estimated is given by Eq. (2) and (3) as follows:

$$W_i = \beta_0 + \beta_1 S_i + \beta_2 X_i + \beta_3 \widehat{P}_i + (\alpha_i + \epsilon_i) \qquad (2)$$

where $\widehat{P}_i$ is the inverse Mills ratio, a transformation of the predicted individual probability of using contacts to find a job from the following selection equation:

$$P_i = \gamma_0 + \gamma_1 Z_i + \gamma_2 X_i + \mu_i \qquad (3)$$

where Z_i is a vector of the factor that explains the choice of job seeking method but should not be included in Eq. (2). In other words, Z_i is the exclusion restriction needed to practically distinguish the sample selection from a mis-specified model.

In the GDR study, I use the dummies denoting whether or not the respondent has siblings and whether or not the respondent has married to serve as the exclusion restrictions. In the China study, marital status as the only exclusion restriction is

imposed because of the family-plan policy implemented in China from middle 1970s. The justifications of using these two variables as the exclusion restrictions are twofold: having siblings or being married leads to a larger pool of potential job helpers, and thus predicts a higher possibility of using contacts to find jobs[71]. Meanwhile, it is reasonable to assume that having siblings or being married does not directly affect the level of wages. Although the validity of exclusion restrictions can hardly be statistically tested for, it is more convincing to use them compared to using the variable of experience which were used in previous studies by Marsden and Hurlbert (1988) and Lin (2003).

The procedure of the Heckit estimation is as follows: in the first step, I use all observations (in this study the whole sample including both contact-users and non-users) to estimate a logit selection model (in this study the model of the choice of job searching method). Then I compute the inverse Mills ratio (or hazard ratio, the estimated probability of not using contacts) for each observation. In the second step, I use the selected sample (in this study those who used contacts) and fit the substantive/outcome model (i.e. the wage model) where the inverse Mills ratio is also included as a covariate to control for the unobserved selection factor.

It should be noted that there might be other sources of endogeneity such as

71 The dummy of having siblings or not has been used as an instrumental variable to identify neighbourhood effects on individual outcomes (see e.g. Bentolila, Michelacci, and Suarez 2010)

ability bias that cannot be corrected for using the Heckit model. To systematically get around the endogeneity problem, based on the Heckit models of the two waves of data, I fit the first-difference model in which the inverse Mills ratio is also incorporated. The first-difference model can be seen as the simplest class of the fixed-effects model, which estimates changes in the dependent variable over time as a function of changes in the independent variables. As a result, the constant effects of time-fixed variables are ruled out from the model. In particular, the first-different model estimates the effect of social capital by comparing the outcome that an individual obtained with a certain level of social capital versus the outcome that the same individual received given another level of social capital. To illustrate this, I first write for two separate cross-sectional wage equations of two different years, namely year *t* and year *(t-1)*, as:

$$W_{it} = \beta_{0t} + \beta_1 S_{it} + \beta_2 X_{it} + \beta_3 \hat{P}_{it} + \alpha_{it} + \epsilon_{it} \quad (4)$$

$$W_{i(t-1)} = \beta_{0(t-1)} + \beta_1 S_{i(t-1)} + \beta_2 X_{i(t-1)} + \beta_3 \hat{P}_{i(t-1)} + \alpha_{i(t-1)} + \epsilon_{i(t-1)} \quad (5)$$

Then I subtract Eq. (5) from Eq. (4) and obtain:

$$W_{it} - W_{i(t-1)} =$$

$$\left(\beta_{0t} - \beta_{0(t-1)}\right) + \beta_1 (S_{it} - S_{i(t-1)}) + \beta_2 (X_{it} - X_{i(t-1)}) + \beta_3 \left(\hat{P}_{it} - \hat{P}_{i(t-1)}\right) + (\alpha_{it} - \alpha_{i(t-1)}) + (\epsilon_{it} - \epsilon_{i(t-1)})$$

(6)

Since I assume that individual unobservable α_i is time-invariant, that is $\alpha_{it} = \alpha_{i(t-1)}$, Eq. (6) ends up as:

$$\Delta W_i = \alpha_0 + \beta_1 \Delta S_i + \beta_2 \Delta X_i + \beta_3 \Delta \widehat{P}_i + \Delta \epsilon_i \qquad (7)$$

where "Δ" denotes the change from t to (t-1). Eq. (7) is thus a first-difference model, predicting changes of outcomes as a function of changes of the independent variables. Since the first-difference models estimate longitudinal models of multiple job searches for the same respondent, I only use individuals who used contacts to find a job more than once during the two time periods. Furthermore, another obvious but crucial condition for estimating Eq. (7), is that S_i must have some variation from time point t to (t-1).

4. EVIDENCE FROM THE GDR

The GDR data collected by Völker and Flap (1999) provides us with an opportunity to do both Heckit and first-difference analyses. The GDR data was collected in 1992, 1993 and 1994 in two cities, Leipzig and Dresden. In this chapter, I focus on the data referring to 1992 and 1993, as the two surveys asked retrospective questions about job and social networks under the socialist period (the first job and job in 1989). Using the unique data set, Völker and Flap (1999) regress the relevant job outcomes (i.e. the occupational prestige of the first job, the occupational prestige of the job in 1989 and wage in 1989) on the contacts' prestige and a set of individual controls (i.e. father's education, the occupational prestige of the father, the education level of the respondents and their sex). To

replicate their study I use the same set of variables except that I include the age variable. The reason for incorporating the age variable is that it can proxy experience, which predicts individual outcomes and increases between the first job and the job held in 1989.

Table 1 presents the selective statistics for two different groups of respondents in two columns: contacts-user for the 1989 job, and contacts-user for both the first and 1989 jobs. The OLS model is to be fitted on the contact-users for the 1989 job, while the first-difference model is to be fitted on the contact-users for both jobs in the second column. The sample size in the first column (194) is exactly the same as that of Völker and Flap (1999). To fit the first-difference models, I drop from the sample those who never used contacts and those who used them only once. There are 116 workers using contacts for their first job, with 194 workers using contacts for their 1989 job. The overlap between them is 66. Thus, the final useful sample size in the second column is 66 workers and the observations for fitting the first difference model are 132. Considering the sample size is too small, for those who merely used contacts for the 1989 jobs, I will recode the "missing value" of the contacts' prestige for their first jobs as zero in the robustness check in order to get more observations.

Table 1: Selected Descriptive Statistics of the GDR Data

	Contact-Users for 1989 job		Contact-Users both first and 1989 jobs	
	N=194		N=132	
Individual attributes	***Mean***	***S.D.***	***Mean***	***S.D.***
Age	38.49	8.02	37.70	8.40
Contacts' Prestige (SEI first job)	--	--	50.13	13.39
Contacts' Prestige (SEI job 1989)	54.02	14.54	54.94	14.37
Occupational Prestige (SEI first job)	48.05	13.03	48.98	13.27
Occupational Prestige (SEI job 1989)	49.45	14.26	50.83	14.69
Occupational Prestige of father (SEI)	44.43	11.83	46.12	11.95
	Freq.	**Per cent**	**Freq.**	**Per cent**
Female	107	55.15%	41	62.12%
Education 1 (father)	25	12.89%	4	6.06%
Education 2 (father)	38	19.59%	11	16.67%
Education 3 (father)	38	19.59%	17	25.76%
Education 4 (father)	55	28.35%	19	28.79%
Education 5 (father)	22	11.34%	10	15.15%
Education 6 (father)	16	8.25%	5	7.58%
Education 1	9	4.64%	2	3.03%
Education 2	24	12.37%	6	9.09%
Education 3	55	28.35%	20	30.30%
Education 4	13	6.70%	3	4.55%
Education 5	35	18.04%	13	19.70%
Education 6	58	29.90%	22	33.33%

Notes: (1) occupational prestige is measured by the Treimen occupational prestige scale for the German version developed with the National Bureau for Statistics.

(2) Education 1 = finished school before 8th grade Education 2 = high school until 8th grade Education 3 = high school until 10th grade Education 4 = high school (university qualification) Education 5 = technical college Education 6 = finished university

The results are presented in Table 2. As mentioned above, although Völker and Flap (1999) collected a two-period panel data (first job and job in 1989), they actually fitted two OLS models respectively for the two time periods and report that there is no evidence supporting the direct effects of using contacts. I thus firstly replicate their study (see column 1), and then proceed to extend their study by fitting the Heckit model to correct for sample selection (see column 2). Before I fit the first difference model to alleviate the remaining endogeneity bias on a smaller sample (see column 5), I also fit the OLS and Heckit model on the same sample for comparison purpose (see columns 3 and 4).

As can be seen from column 1 of Table 2, explanatory variables that have significant effects are the contacts' prestige, educational attainment, and the prestige of the first job. Using the coefficient on contacts' prestige, we can see that getting a job from a contact with a one unit higher occupational prestige score is associated with a .19 higher occupational prestige score of the job held in 1989, everything else being equal. This pattern is very similar to what has been found in the pioneer research by Völker and Flap (1999)[72]. When it comes to the results obtained from the Heckit model in column 2, one can see that it (.17) is slightly lower than that from the OLS model. However, the magnitude is very close and the effect is significant given a .05 alpha level. Notice that the Heckit estimate is lower than obtained from the OLS model. Since Chapter 5 on the same

[72] The coefficient on contacts' prestige estimated by Völker and Flap (1999) is 0.17.

data suggests that workers of higher unobserved ability in the GDR tend to use contacts to find jobs, we could explain this by arguing that workers of a higher ability benefit more from the embedded resources. This may be a case of the Mathew Effect playing a part in mobilising social resources.

Turn now to the estimation results from the first-difference model following the same workers over time. One way to think about this method is that we predict deviations of status from the respondents' average status as a function of deviations from the respondents' average of a set of time-variant independent variables. Consequently, this model estimates the effect of contacts' prestige on status for those who used contacts for both previous and current jobs. However, this substantially decreases the sample size from 194 to 66 workers. Furthermore, it is true that the 66 workers who used contacts for both the first and the 1989 job are more likely to share some unobserved individual attributes. To be careful, I first of all fitted the OLS model on the 66 workers used in the first-difference analysis. As can be seen in column 3, although the estimates change a lot, the general pattern of the estimated effects of the set of independent variables remains stable. I then proceed to fit the Heckit model so as to deal with the sample selection and present the results in column 4. It is easy to see that the estimated effect of social capital in column 4 is very close to that in column 2, suggesting that the Heckit model does indeed alleviate the selection bias due to the shared unobservables of the 66 workers.

Table 2: OLS, Heckit and First-Difference Models on GDR Data (O ccupational prestige as the dependent variable)

Independent Variables	OLS Contact-Users for 1989 job	Heckit Contact-Users for 1989 job	OLS Contact-Users for both jobs	Heckit Contact-Users for both jobs	Heckit-FD Contact-Users for both jobs
Prestige of contact	.192***	.174***	.331**	.172**	.209**
	(.069)	(0.063)	(0.154)	(0.075)	(0.089)
Sex (male=0)	.87	1.30	.10	.66	--
	(1.763)	(1.869)	(3.418)	(2.371)	--
Education	1.87***	1.94***	2.2**	1.54**	--
	(0.689)	(0.792)	(0.101)	(0.772)	--
Father's education	-.58	-.21	-1.18	.55	--
	(0.471)	(0.677)	(1.401)	(.183)	--
Prestige of father's job	.057	.003	.052	.075	--
	(0.079)	(0.085)	(0.164)	(0.104)	--
Prestige of the first job	.40***	.44***	.47***	.53***	--
	(0.089)	(0.094)	(0.099)	(0.110)	--
Age	.17	.37	.58	.36	-1.92
	(1.199)	(1.286)	(1.999)	(1.287)	(2.174)
Age-squared	-.0002	-.0027	-.0073	-.0037	.024
	(0.015)	(0.016)	(0.024)	(0.016)	(0.028)
Inverse Mills ratio	--	.85***	--	.87***	-1.45
	--	(0.084)	--	(0.094)	(1.310)
_cons		-9.89	2.77	4.42	3.59
		(24.89)	(3.651)	(27.89)	(11.809)
N of cases	194	194/476	66	66/194	66
N of observations	194	194/476	66	66/194	132

Note: i) Standard errors are in parentheses and robust to heteroscedasticy;
ii) * p < =0.1 ** p <= 0.05, *** =p < 0.01
iii) FD=First Difference

i

The first-difference estimates in column 5 are even more meaningful because any time-fixed unobserved individual variables will drop out[73]. Comparing column 3 and column 5, one can find that the first-difference estimates of the effect of contacts' prestige decreases from .33 to 0.21, yet is still significant at the 0.05 alpha level. That is, without controlling for the time-invariant unobservables and sample-selection, the effects of the contacts' prestige are upwardly biased. Comparing column 4 and column 5, we find that the first difference estimate is slightly larger than the Heckit estimate. One possible explanation is that unobserved ability is negatively correlated with social capital. The reason is that, since it is reasonable to expect that the unobserved ability should be positively correlated with individual outcomes, the estimation will be downwardly biased if the contact's status is negatively correlated with the unobserved ability. Put differently, job seekers of less ability may have adopted a compensation strategy and turn to more resourceful contacts to get help to find a job. A similar phenomenon can be found in Chapter 4 where I report this type of heterophilous selection on educational social capital as "unfruitful" resources during the making of core discussion networks among Chinese in 2000s. Since job status of contacts is more likely to be an indicator of "fruitful" resources embedded in social ties, the heterophilous selection here could be explained by the inequality of social resources under state socialism in the GDR in early 1990s. However, if Treimen occupational prestige scale for the German version used in the GDR

[73] I also incorporate the reverse Mills ratio (the selection factor) of the Heckit model into the first difference model.

study does not reflect the genuine rating of status/resources in this country under state socialism, contacts' prestige measured by Treimen scale could be also regarded as a type of "unfruitful" resources which accounts for the heterophilous selection.

As mentioned before, to ensure thoroughness, I do a robustness check on a larger sample. That is, for those who did not use contacts for the first jobs but used contacts for their 1989 jobs, the used social resources for getting the first job are regarded as zero. This increases the observations from 132 workers to 388. Replacing the missing value with zero makes sense especially for this group of people. In contrast, for those who used contacts for their first jobs but not for the 1989 jobs, things are more complicated. This is because it is possible that the resource of the contact for the first job still matters in finding the 1989 job. If so, it is problematic to assume the missing value to be zero for this group of people. Although not reported here, the estimates obtained from the increased samples are very similar to what is reported in the third column in Table 2. The estimated effect obtained from the analysis of the 388 observations is .19 and is statistically significant.

5. EVIDENCE FROM CHINA

Exploiting multiple identification strategies, I have shown that there is a causal

link between the used social capital and the attained status of workers in the GDR. Can this pattern be observed in other socialist countries and one with such a different culture as China? There has been a large body of literature on the use of social networks in the labour market in China. Bian (1997) reports that Chinese people relied more on strong ties to find jobs under the job assigning system before the reform and opening-up. Bian and Ang (1997) argue that the occupational prestige of the job helper positively predicts the attained status of the job changer. In a more recent study, Lin (2003) finds that the contacts' supervising unit rank affects the job seekers' supervising unit rank in a positive way. That being said, however, the endogeneity problem has not been adequately addressed in these earlier studies. In this section, I will probe whether the used social capital, or the status of the job helper, influences the wage of the job seeker.

The data which will be used in the present analysis is from the 2002 CHIPS which is also used in Chapter 5, containing a sample of 20,632 urban residents from 22 provinces. The questionnaires gathered detailed information on demographics, recent employment history and wages of the current and previous jobs of urban respondents. This offers the opportunity to fit the fix-effects model on panel data to probe the causal effect of the social resource used to get a job.

The questionnaire used in CHIPS 2002 asked respondents to provide detailed information regarding the respondents' current and previous jobs. In CHIPS 2002,

the respondents were asked the question: "Have you changed jobs in the last three years?"[74] This allows me to compare the different results between different levels of social resources as used by the same respondent to find jobs within a three-year period. Notice that, in estimating the first-difference model, those who did not change their job from 2000 to 2002 will drop out.[75] In addition, respondents who are younger than 16 or older than 60 are also not examined. Considering the potential outliers in terms of wages, I also dropped 236 respondents who reported zero wages.[76] To fit the first-difference model, those who never used contacts and those who used them only once also drop out of the sample[77]. Finally, I get 764 workers who had used contacts for their current jobs, while 130 workers who had used contacts for their previous jobs. Importantly, the overlap between them is 71. Thus, the final useful sample for carrying out the first-difference analysis is 142 observations. Again, in the robustness check, I will not remove those who merely used contacts for their current job and therefore get a sample of 424 observations.

74 This refers to the period from 2000 to 2002.

75 Hereafter "previous job" refers to the job before the current job after 1999, although many non-changers (1999–2001) did have previous jobs before 1999.

76 In robustness checking, I also dropped observations in the top 1% and bottom 1 % from the sample. The results are very stable.

77 As for the job search method, the exact phrasing of the question reads: "How did you get your current job?" and "How did you get your previous job?". Ten possible answers are offered: 1) assigned by government; 2) inherited it; 3) through an open examination; 4) through an employment agency; 5) through the newspaper; 6) through a referral by someone you or your family knows; 7) found it on your own; 8) started your own business; 9) not working currently; 10) other. Only one answer can be chosen. If the respondent chooses the sixth answer, the respondent would be classified into the group using contacts.

Table 3: Selected Descriptive Statistics of CHIPS 2002 Data

	Contact-Users for current job		Contact-Users for both jobs (within last 3 years)	
	N=764		N=142	
Individual attributes	***Mean***	***S.D.***	***Mean***	***S.D.***
Age	39.47	8.38	36.17	8.69
Years of schooling	10.33	2.81	10.66	2.62
Annual wage of prior job (RMB)	--	--	8420	8939
Annual wage of current job (RMB)	10995	9490	9120	7359
Experiences for prior job	--	--	18.86	9.64
Experiences for current job	23.15	9.37	19.50	9.82
	Freq.	**Per cent**	**Freq.**	**Per cent**
Female	402	52.62%	36	50.70%
Contact's CCP membership (prior job)	--	--	19	26.76%
Contact's CCP membership (current job)	356	46.60%	20	28.17%
Work unit sector (current job)				
Governments/institutes	146	19.11%	8	11.27%
State-owned enterprises	152	19.90%	10	14.08%
Collectives	61	7.98%	3	4.23%
Joint ventures	28	3.66%	4	5.63%
Private/self-employed	377	49.35%	46	64.79%
Work unit sector (prior job)				
Governments /institutes	--	--	6	8.45%
State-owned enterprises	--	--	11	15.49%
Collectives	--	--	4	5.63%
Joint ventures	--	--	2	2.82%
Private/self-employed	--	--	48	67.61%

In the present analysis, I try to understand the contacts' status in the context of social transition in China and measure it with contacts' membership of the CCP[78]. It has been broadly discussed that CCP membership reflecting political loyalty plays a significant role in the socialist stratification process (Gerber 2000; Wong 1996), described by Walder et al. (2000: 192) as a system in which: "[the party] operated as a conventional political machine that rewarded loyalty by allocating desirable jobs". Although the path-dependence argument maintains that the pre-existing political capital would take the new form of socioeconomic capital and would therefore ensure that the individual was able to maintain their advantageous position (Rona-Tas and Guseva 2001; Bian and Logan 1996; Shu and Bian 2003; Walder 2003), the political loyalty under a market economy in the making can be speculated to play less of a role than it did in Mao's era (see e.g. Nee 1989, 1991, 1996). As Wu and Xie (2003) point out, in the debate controversies are mainly focused on to what extent redistributive power affects individual outcomes. What motivates the present China study is therefore twofold: first, to assess whether contact's CCP membership matters for job-seeking in the presence of sample selection and omitted variables; second, to explore whether CCP membership can serve as a strong indicator of valuable resources or weak one. In CHIPS 2002, the exact phrasing of the question of the political status of contacts reads: "If you found a job through a relative or a friend,

78 As mentioned in Chapter 5, we do not have a well-established and recognized index or scale of occupational prestige for China under social transition.

was he/she a communist party member?" The selective statistics are presented in Table 3.

Table 4 below presents the estimation results obtained from the OLS model, Heckit model and the first-difference model. As is shown in column 1 , the results from the OLS model demonstrate that, among those who used contacts to find their current jobs, others being equal, contacts-users getting help from a contact who is a CCP member is associated with a 12.3 per cent higher annual wage in comparison with using a contact who was not a CCP member. As for other control variables, gender, educational attainment, work unit sector and location all have an effect on wages. When it comes to the estimates from the Heckit model on the same full sample(column 2), the results reported in the second column show that, controlling for the unobserved individual characteristics which determined whether or not the respondent used contacts to find jobs, the effect of the political status of the contact remain significant and has even increased from .12 to .13. Based on the findings in Chapter 5 that less able Chinese tend to use contacts, a plausible explanation is that less able workers in China benefit less from using contacts. Just as I reported in the Heckit analysis based on the GDR data, this also reflects the Mathew Effects in the mobilisation of embedded resources.

Table 4: OLS, Heckit and First-Difference Models on CHIPS 2002 Data
(Log-wages as the Dependent Variable)

Independent Variables	**OLS** Contact-Users for current jobs	**Heckit** Contact-Users for current jobs	**OLS** Contact-Users for both jobs	**Heckit** Contact-Users for both jobs	**Heckit-FD** Contact-Users for both jobs
Contact being CCP	.123***	.131***	.162	.117**	.181**
	(0.046)	(0.046)	(0.199)	(0.049)	(0.020)
Female	-..21***	-.138***	-.233	-.200***	--
	(0.044)	(0.050)	(0.169)	(0.047)	--
Education	.067***	.028**	0.060**	0.070***	--
	(0.0093)	(0.012)	(0.025)	(0.001)	--
Experience	.0086	.002	.005	.005	.231*
	(.0094)	(.0011)	(0.032)	(0.011)	(0.127)
Experience-squared	-.0006	-.00002	-.0011	.0001	-.005*
	(.0005)	(.002)	(.007)	(.0002)	(0.003)
Inverse Mills ratio	--	.76***	--	.60**	.80
	--	(0.043)	--	(0.020)	(0.65)
Work unit sector					
State-owned	-.146**	-.193**	-.376	-.122*	-.28
	(0.068)	(0.076)	(0.364)	(0.073)	(0.308)
Collectives	-.28**	-.16	.135	-.299**	-.86**
	(0.094)	(0.103)	(0.246)	(0.095)	(0.29)
Joint ventures	.069	.39**	-.0004	.054	.084
	(0.106)	(0.155)	(0..33)	(0.139)	(0.382)
Private/self-employed	-.204***	-.155**	-.326	-.189***	-.043
	(0.063)	(0.07)	(0.288)	(0.069)	(0.325)
Provinces Dummy	YES***	YES***	YES***	YES***	--
_cons	9.12***	8.41***	8.98***	9.11***	.47*
	(0.217)	(0.266)	(0.770)	(0.241)	(0.238)
N of observations	764	764/9366	71	71/764	142

Note: i) Standard errors are in parentheses and robust to heteroscedasticy;
ii) * p < =0.1 ** p <= 0.05, *** =p < 0.01

Before proceeding to the first-difference analysis, I fit the OLS model on those who used contacts for both their previous and current jobs. This indeed dramatically shrinks the sample size from 764 to 71. As can be seen in column 3 in Table 4 above, the significant effect of the political status of contacts disappears when I use the decreased sample. Obviously the sample is very small and most likely to be biased because respondents using contacts twice within three years may share certain unobserved individual attributes. Thus I also fit the Heckit two-stage model to alleviate the sample selection bias. As is shown in column 5, the role of social capital is significant, which is obviously different from the result that I have obtained from the standard OLS model in column 4. This finding suggests that the effect of contacts' political status is dramatically downwardly biased if the sample selection is not taken into account.

If we turn now to the estimates obtained from the first-difference model in column 5, we can see that the estimated effect of getting help from a contact who is a CCP member is 0.181 and can be significantly distinguished from zero given a 0.05 alpha level. The estimated effect is thus slightly higher than that obtained from the Heckit model. The findings are thus twofold: first, contacts' political status pays off in the job matching process in China controlling for time-fixed confounders; second, as in the GDR analysis, the first difference estimate is larger than the Heckit estimate. This implies that individuals of less earning ability adopt a compensation strategy to turn to job contacts who are in possession of more resources. Although counter-intuitive to an extent, this pattern has been

found in Chapter 4 when I addressed whether the average of close friends mattered for individual outcomes. The interpretation here is, CCP membership as a specific resource of political status brings about benefits, but it may have become less "fruitful" in early 2000s as compared to the pre-reform era. The implication of the compromised role of the CCP membership in early 2000s can actually be supported by both sides of the debate on China's social transition which is mentioned above.

Before closing this section on Chinese data, I carried out a robustness check on a larger sample which includes those who did not use contacts for the previous jobs but used contacts for their current job. This increases the observations from 142 to 424 (212 respondents). Although not reported here, the estimates obtained from the increased samples are very similar to those reported in the fourth column in Table 3. The estimated effect obtained from the analysis on 424 observations is .151 and statistically significant. In addition, I include the interaction between the political status of contacts and the strength of the tie between job seekers and job helpers. However, this does not affect the results reported in Table 4. Finally, since not all party members have the chance to obtain an elite managerial position, I use being a CCP cadre rather than a CCP member to serve as the explanatory variable. Although not reported here, the results are basically the same.

6. CONCLUSION

While it is theoretically intriguing and intuitive that contacts matter for individual outcomes, most previous studies giving supportive evidence (Bridges and Villemez 1986; Marsden and Hurlbert 1988; Völker and Flap 1999; see Lin 2008 for a review) have not adequately addressed the endogeneity problem. This is especially problematic in the research setting of the job matching process because the use of contacts to find a job is very likely to be selective and unobservable associated with both social capital and individual outcomes can bias the estimation. Consequently, naively fitting an OLS model to examine the relationship between contacts' resources and the respondent's status or wages tends to bring about biased estimates.

In this chapter, I address the effects of contacts' occupational prestige on a job seeker's occupational prestige in the GDR, and analyse whether a contact that is a CCP member brings about a better job as compared to a contact who is not a CCP member in contemporary China. Paying very close attention to the potential endogeneity problem, I compare the estimations from the standard OLS model, the Heckman two-stage model and the first-difference model. The results indicate that the occupational prestige of a contact matters for the status attainment in the GDR, and the CCP membership of a contact influences the level

of individual wage in China. The study also reveals the Mathew Effects in using social capital：an individual of higher unobserved ability tends to benefit more from using contacts in both the GDR and China. Furthermore, individuals of poor unobserved ability may have used a compensation strategy to look for more "resourceful" job help to find jobs under state socialism. However, resources that are available for them are often "unfruitful".

The contribution of this chapter to the accumulated sociological literature of used social capital is the use of a multi-method to alleviate potential estimation biases led by unobserved individual ability and selective social ties. Using the Heckit model and the first-difference model to re-analyse the GDR data, the present study confirms the findings in the previous study by Völker and Flap (1999) that social resources matter in status attainment in the GDR. Applying the method to data collected from China, moreover, this study offers more evidence in favour of the labour market role of contacts under state socialism. In addition, comparing results obtained from difference models sheds more light on our understanding of how people select contacts and how contacts influence job outcomes. In general, we can state that so far we can confidently argue that contacts matter for job outcomes at least in these two countries under state socialism. To be more cautious, given the small sample size used in the first difference analysis, larger data sets and data from more countries are needed in future studies.

Chapter 7

Conclusion

This thesis analyses the role of social capital in the labour market, which is defined as the characteristics of one's social network. It is theoretically intriguing and intuitionally appealing to think that the resources embedded in ones' social network can positively influence individual socioeconomic outcomes such as employment, occupational prestige, wage or income. However, as Chapter 2 has shown, to confirm the causal link between characteristics of social networks and individual outcomes is difficult. This is because there are various sources of endogeneity bias when estimating models with observational data. To identify causality under the counterfactual framework, four empirical chapters (chapters 3-6) in this thesis employ the instrumental variable method, the Heckman two-stage method and the first difference method to alleviate the estimation biases. This thesis contributes to this line of research by using a constellation of identification strategies to test for the causal effects of social capital in a systematic way.

1. EMPIRICAL FINDINGS

In Chapter 3, I analysed the network effects among rural-to-urban migrants in

China based on data from 2002 Chinese Household Income Project Survey (CHIPS 2002). Since it is well established that Chinese rural-to-urban migrants return home frequently in a year, the job-related information at different destinations can be effectively transmitted within the origin village. Consequently, in a given year, the total outflow of migrants from one's origin village is very likely to positively influence his/her wage at the destination by offering useful job leads. To test for this hypothesis, one must take into account the selection process in the migration decision and other confounders. I thus take advantage of the instrumental variable (IV) method combined with the Heckman two-stage method to estimate the network effects among migrants. Using natural disaster as the IV, this chapter reveals that the size of migrant network of origin village significantly improves the wages of its member migrants. Furthermore, in comparison with the standard Heckit estimates, the OLS estimates are downward-biased, suggesting that farmers who are of higher earning-ability tend to migrate, if we believe more able migrants benefit less from the origin-based networks. Compared to the IV-Heckit estimates, however, the standard Heckit estimates are downward-biased. This can be interpreted under the Local Average Treatment Effects framework. Specifically, less able people may be more responsive to natural disasters and the IV estimate is larger because it mainly captures network effects among less able migrants who benefit more from networks based on original villages. On balance, therefore, these findings strongly support the positive role of migrant networks on wages, and reveal that

naively fitting a standard OLS model leads to considerable estimation bias.

In Chapter 4, my aim was to identify the effects that network resources have on individuals' socioeconomic outcomes in the presence selectivity in urban China. Using the 2003 Chinese General Social Survey data, I assess how friends' educational attainment affects one's labour force participation and income. In the empirical analysis, a major concern is that, since social networks are often formed in a non-random way, the estimated effects of network resources in many previous studies are very likely to be biased and just reflect the presence of social homophily in the process of friendship formation. Considering this, I instrumented for each respondent's observed network resource with the average network resource of a set of observationally identical respondents. The empirical results demonstrate that the IV estimates of the network effects on employment and income are substantially larger than those obtained from single equation models (the Probit and OLS models). The significant and larger IV estimates imply that individuals with poor unobservables in terms of employment or income (e.g. less earning-ability, disadvantaged positions in labour markets etc.) might have adopted a compensation strategy to select friends with higher educational attainment. This is consistent with Lin's (2008) finding that, when the resources embedded in homophily-based networks are not sufficient for a specific goal, individuals may reach beyond the inner layer of their social networks to access heterophilous resources. In general, this chapter gives further

weight to Mouw's (2003, 2006) call for taking seriously the selection problem when trying to identify the causal effects of network resources. Given that the IV is based on the working assumption that relevant controls in the model are exogenous, the estimates from the IV models are meaningful not because they are more precise or unbiased but because the IV method can address the additional bias led by self-selection.

In Chapter 5, I revisited the old proposition that using contacts to find a job pays off in comparison to using other channels. The empirical results in previous sociological studies are mixed and certainly questionable, given that using or not using contacts is not a random choice. To adjudicate between the argument that using contacts brings about advantages and the argument that using contacts does not pay off, I conducted an empirical study analysing data from two countries under state socialism: East Germany (GDR) and China. More specifically, I replicated and extended Völker and Flap's study (1999) on the GDR data. Using the same model, I also analysed the data from the CHIPS 2002. Both the GDR and Chinese results are consistent with previous findings showing no consistent evidence supporting the benefits of the use of contacts by itself as compared to the use of other channels to find jobs. However, the GDR study reveals that the effect of using contacts on status attainment is upwardly biased if one does not take into account potential selection, while the Chinese study shows that the standard OLS estimate of wage effect is downwardly biased. Based on this study,

the picture is that the GDR sees positive selection while China sees negative selection. A potential explanation is that Chinese workers of lower ability tend to use contacts to find jobs, while in the GDR the opposite may be the case.

In Chapter 6, using the same two datasets and the same analytical strategy as in Chapter 5, I estimated the causal effects that the resources embedded in social structure have on individual outcomes in the labour market under state socialism. To deal with the potential sample selection, I exploited the Heckman two-stage model to allow for an unrepresentative sample. To correct for general omitted variable biases, I followed the same respondents over time to see whether any significant difference develops between the time when they used contacts for the previous jobs and used contacts for the current jobs in the surveying year. As with the method employed in Chapter 5, in Chapter 6 I replicated and extended the study by Völker and Flap (1999) to identify the causal effects of social capital in the GDR. I then used the same identification strategy to analyze the Chinese data to obtain more evidence under state socialism. The results from the GDR data confirm the major findings of previous studies: contacts' occupational prestige matters for workers' occupational prestige. The evidence from China also shows that there is a causal relationship between the contacts' membership of the Chinese Communist Party – political status as a specific type of social capital – and respondents' wages. Interestingly, this chapter reveals that in both China and the GDR, individuals of less earning ability adopt a compensation

strategy to turn to job contacts in possession of more resource. This finding about the compensation strategy is consistent with the results in Chapter 4.

2. LIMITATIONS

I realise that the findings in this thesis have two major limitations.

First, the findings are subject to the validity of the working assumptions needed for fitting relevant advanced models. One stylised fact of social network data is the importance of unobserved heterogeneity across individuals. To deal with potential estimation bias led by various confounders, I used a constellation of identification strategies to estimate models. However, the methods used in this thesis are subject to certain working assumptions. For instance, when fitting the first-difference model in chapters 5 and 6, a crucial assumption is that individual heterogeneity is time-fixed. In the GDR study, given the time span between the first job and the job in 1989, I need to estimate under a strong assumption that during the years the unobserved heterogeneity had not varied. In the study on network effects among rural-to-urban migrants (Chapter 3), the estimated effects of the migration networks based on origin village are subject to the validity of the assumption that a natural disaster affecting the village is totally exogenous. In Chapter 4, I even constructed the IV based on the assumed exogeneity of other controls, which must be subject to criticism. However, empirical analyses under

these assumptions are the most the data can tell us without additional unverifiable assumptions. In addition, presenting and comparing estimates from advanced models and from standard models are useful because they together shed more lights on our understanding about the sources of endogeneity.

Second, some studies in this thesis suffer from very limited data. For example, the sample size of the study on the treatment effects of using contacts as compared to not using contacts (Chapter 5) is merely around 430 for the GDR case and 750 for the China case. This questions the reliability of the Heckman estimates, as the Heckit method depends on having a large sample size. In Chapter 6, estimating the effects of the used social resources, fitting the first-difference model leads to a very small sample size of less than one hundred.

3. EXTENSION

First, more attention should be paid to the mechanism of selection in the process of forming and using social ties. Studies focusing on the causal effects of social networks (Moue 2003, 2006) strongly suggest that the causal relationship between social networks and individual outcomes may be over-estimated due to the presence of selection. However, this thesis demonstrates that the effects of social networks can be under-estimated without taking into account the selection. For example, in Chapter 4, I find people with unobserved less earning

ability have friendship networks with people with a better level of education, since the IV estimates are larger than standard OLS estimates. Chapter 6 also reveals that the first-difference estimates are larger than standard OLS estimates, suggesting that individuals with less earning ability turn to job contacts in possession of more resources in both China and the GDR. To sum up, these findings strongly suggest the presence of negative selection based on social heterophily rather than social homophily. However, this thesis has only touched on the mechanism of this type of negative selection. More academic attention is thus called for in future studies in this line of research.

Second, high-quality data in relation to social networks are needed in order to identify the causality. Sometimes, network analysts have to deal with very limited data, often cross-sectional and small in size, something that can be entailed by the features of social networks. For example, it may be very difficult to collect panel data because the members and structure of the egocentric social networks can be very stable for a very long time. However, to overcome the threat of endogenity, the sociologist must rely on datasets with more information, which can be of help to identify models. Three types of data can be very useful for identifying the causal effects of networks: (1) time series data, which allow us to compare the variation in the same respondents over time. These data are useful as we can exploit them to rule out the time-fixed heterogeneity; (2) data from more extensive networks. They can be beneficial as the overlapped section

among people's egocentric networks can be instrumented under certain circumstances. For example, the occupational prestige of the friends of the friends of a respondent must be related to the occupational prestige of his/her friends (explanatory variable of interest), but does not directly influence the social outcomes of the respondents (dependent variable of interest); (3) genuine exogenous variables such as natural events. At first glance, such data may not be related to the hypotheses of interest, but in fact their usefulness arises from this "irrelevance". In econometrics literature, information ranging from earthquakes to volcano eruption has been exploited as the exclusion restrictions.

Third, more replication studies based on shared data are called for. I echo Jeremy Freese's argument (2007a, 2007b) that quantitative sociology would benefit from more transparency and replication. This is especially true because some data on which a previous study was based on can provide us with new information needed to deal with the endogeneity problem, which was often ignored by previous studies. In this thesis, I conduct two replication studies on the GDR data collected and analysed by Völker and Flap (1999). A comparison based on a replication study helps us to understand why the endogeneity problem is important and how the confounders bias the naive estimates. The aim of conducting a replication study is by no means to pick on existing literature, but rather to increase the credibility of results and improve the quality of data analysis. It is for this reason that such extensions are called for.

References

Aaronson, Daniel. 1998. “Using Sibling Data to Estimate the Impact of Neighborhoods on Children’s Educational Outcomes.” *Journal of Human Resources* 33 (4): 915-46

Alexander, David. 2000. *Confronting Catastrophe*. Oxford University Press: New York, NY.

Ammermüller, Andreas, and Jörn-Steffen Pischke. 2006. “Peer Effects in European Primary Schools：Evidence from PIRLS.” *ZEW Discussion Paper* No. 06-027, Mannheim.

Angrist, Joshua D., Guido W Imbens, and D.B. Rubin. 1996. “Identification of Causal Effects Using Instrumental Variables.” *Journal of the American Statistical Association* 91: 444-55.

Angrist，Joshua D., and Alan B. Krueger. 1999. “Empirical Strategies in Labor Economics.” In Orley Ashenfelter and David Card, (eds.) *Handbook of Labor Economics IIIA*. Amsterdam: North-Holland.

Angrist，Joshua D., and Victor Lavy 1999. “Using Maimonides' Rule to Estimate the Effect of Class Size on Student Achievement.” *Quarterly Journal of Economics* 114(2):533-75.

Araujo, Caridad, Alain de Janvry, and Elisabeth Sadoulet. 2010. "Peer Effects in Employment: Results from Mexico's Poor Rural Communities". *Canadian Journal of Development Studies* Volume 30, number 3-4

Arcidiacono, Peter, and Sean Nicholson. 2005. “Peer Effects in Medical School.” *Journal of Public Economics* 89(2-3):327-50.

Asadullah，Mohammad Niaz, and Nazmul Chaudhury. 2008. “Social Interactions and Student Achievement in A Developing Country: An Instrumental Variables Approach.” *Policy Research Working Paper Series* 4508. The World Bank.

Balderrama, Rafael and Hilario Molina. 2009. “How Good Are Networks for Migrant Job Seekers? Ethnographic Evidence from North Carolina Farm Labor Camps”. *Sociological Inquiry,* 79(2): 190–218.

Banerjee, Biswajit. 1983. “Social Networks in the Migration Process: Empirical Evidence on Chain Migration in India.” *The Journal of Development Areas*. 17: 185-96.

------ 1984. “Information Flow, Expectations and Job Search: Rural-to-Urban Migration Process in India.” *Journal of Development Economics* 15: 239-57.

Bao, Shuming, Örn B. Bodvarsson, Jack W. Hou, and Yaohui Zhao. 2007. "Interprovincial Migration in China: The Effects of Investment and Migrant Networks." *IZA Discussion Paper* 2924.

Bauer, Thomas, Pedro T. Pereira, Michael Vogler, and Klaus F. Zimmermann. 2002. "Portuguese Migrants in the German Labor Market: Selection and Performance." *International Migration Review* 36(2): 467-91.

Bayer，Patrick, and Stephen L. Ross. 2009. "Identifying Individual and Group Effects in the Presence of Sorting: A Neighborhood Effects Application." *NBER Working Paper* 12211. Cambridge，United States：National Bureau of Economic Research.

Bayer, Patrick, Stephen Ross, and Giorgio Topa. 2005. "Place of Work and Place of Residence: Informal Hiring Networks and Labor Market Outcomes." *NBER Working Paper* No. 11019 Cambridge，United States：National Bureau of Economic Research

Bayer，Patrick, Robert McMillan, and Kim Rueben. 2002. " What Drives Racial Segregation? New Evidence Using Census Microdata." Working Papers 02-26

Beaman, Lori A. 2008. "Social Networks and the Dynamics of Labor Market Outcomes : Evidence from refugees," Mimeo Northwestern University

Bentolila, Samuel, Claudio Michelacci, and Javier Suarez. 2010. "Social Contacts and Occupational Choice." *Economica* 77(305): 20–45.

Berk，Richard A., and Subhash C. Ray. 1982. "Selection Biases in Sociological Data." *Social Science Research* 11:352-98

Berk，Richard A. 1983. "An Introduction to Sample Selection Bias in Sociological Data" *American Sociological Review* 48(3): 386-98.

Betts, Julian R., and Andrew Zau. 2004. "Peer Groups and Academic Achievement: Panel Evidence from Administrative Data." Working Paper. University of California, San Diego

Bian, Yanjie. 1994. "Guanxi and the Allocation of Urban Jobs in China". *The China Quarterly* 140:971-99.

---- 1996. "Chinese Occupational Prestige: A Comparative Analysis." *International Sociology* 11(2):161-86.

---- 1997. "Bringing Strong Ties Back in: Indirect Ties, Network Bridges, and Job Searches in China." *American Sociological Review* 62(3):366-85.

----2008, “The Formation of Social Capital among Chinese Urbanites: Theoretical Explanation and Empirical Evidence”. Pp. 81-104. In Nan Lin and Bonnie Erickson (ed.) *Social Capital: An International Research Program*. Oxford and New York: Oxford University Press.

Bian, Yanjie, and John R. Logan. 1996. “Market Transition and the Persistence of Power: The Changing Stratification System in Urban China.” *American Sociological Review* 61(5):739–58.

Bian, Yanjie, and Soon Ang. 1997. “Guanxi Networks and Job Mobility in China and Singapore.” *Social Forces*, 75(3):981-1005.

Blalock, Hubert M., Jr. 1984. “Contextual Effects Models: Theoretical and Methodological Issues.” *Annual. Review of Sociology* 10:353-372.

Blau, Peter. M. 1977. *Inequality and Heterogeneity*. New York: Free Press.

Bobonis, Gustavo J., and Frederico Finan. 2008. “Neighborhood Peer Effects in Secondary School Enrollment Decisions.” Working paper.

Boozer，Michael, and Stephen E. Cacciola. 2001. “Inside the ‘Black Box’ of Project Star：Estimation of Peer Effects Using Experimental.”Working paper.

Boxman, Ed A. W., Paul M. De Graaf, and Hendrik D. Flap. 1991. “The impact of social and human capital on the income attainment of Dutch managers.” *Social Networks* 13:51-73.

Boyd, Monica. 1989. “Family and Personal Networks in International Migration: Recent Developments and New Agendas.” *International Migration Review* 23(3): 638-70.

Bramoullé, Yann, Habiba Djebbari, and Bernard Fortin. 2007. “Identification of Peer Effects through Social Networks.” IZA Discussion Paper 2652.

Bridges William P., and Wayne J. Villemez, 1986, “Informal Hiring and Income in the Labor Market.” *American Sociological Review* 51(4):574-82.

Briggs, Xavier de Souza. 1998. “Brown Kids in White Suburbs: Housing Mobility and the Multiple Faces of Social Capital”. *Housing Policy Debate* 9(1): 177-221.

---- 2002. “Social Capital and Segregation: Race, Connections, and Inequality in America”. Working Paper RWP02-011. Kennedy School of Government, Harvard University.

Brock，William A., and Steven N Durlauf. 2001. “Interactions-based Models.” In J.J. Heckman and E.E. Leamer (ed.) *Handbook of Econometrics*, edition 1, volume 5, chapter 54: Elsevier.

Bushway, Shaun, Brian D. Johnson, and Lee Ann Slocum. 2007. “Is the Magic still There? The Relevance of the Heckman Two-step Correction for Selection Bias in Criminology.” *Journal of Quantitative Criminology* 23(2):151-78.

Burt, Ronald S. 1992. *Structural Holes: The Social Structure of Competition*. Cambridge, MA: Harvard University Press.

Cai, Fang. 1996. “An Economic Analysis for Labor Migration and Mobility.” *China Social Sciences Quarterly,* Spring: 120-135.

Calvó-Armengol, Antoni, and Yannis Ioannides. 2005. “Social Networks in Labor Markets”, in L. Blume and S. Durlauf (eds) The New Palgrave, *A Dictionary of Economics*, Second Edition. London: MacMillan.

Calvó-Armengol, Antoni, and Yves Zenou. 2005. “Job Matching, Social Network and Word-of-Mouth Communication.” *Journal of Urban Economics* 57(3):500-522.

Cameron, A. Colin, and Pravin K. Trivedi. 2005. *Microeconometrics: Methods and Applications.* Cambridge：Cambridge University Press.

Campbell, Karen E., Peter V. Marsden, and Jeanne S. Hurlbert. 1986. “Social Resources and Socioeconomic Status”. *Social Networks* 8:97-116.

Campbell, Karen E. and Rachel A. Rosenfeld. 1985. “Job Search and Job Mobility: Sex and Race Differences.” *Sociology of Work* 3: 147–74.

Card, David, and Alan Krueger. 1996. “School Resources and Student Outcomes: An Overview of the Literature and New Evidence from North and South Carolina.” *Journal of Economic Perspectives* 10(4): 31-50.

Card, David. 1995. “Using Geographic Variation in College Proximity to Estimate the Return to Schooling.” In L.N. Christofides, E.K. Grant, and R. Swidinsky, (eds.) *Aspects of Labor Market Behaviour: Essays in Honour of John Vanderkamp*. Toronto: University of Toronto Press, 1995.

------ 1998. “Education in the Labor Market.” In O. Ashenfelter & D. Card (eds.) *The Handbook of Labor Economics*, North-Holland, Amsterdam.

Carman, Katherine, and Lei Zhang. 2008. “Classroom Peer Effects and Academic Achievement: Evidence from a Chinese Middle School.” Working paper.

Case, Anne C., and Lawrence F. Katz. 1991. "The Company You Keep: The Effects of

Family and Neighborhood on Disadvantaged Youths," NBER Working Paper No. 3705.

Cason，Timothy N., and Vai-Lam Mui.1998. “Social Influence in the Sequential Dictator Game.” *Journal of Mathematical Psychology* 42: 248–65.

Carrington, William J., Enrica Detragiache, and Tara Vishwanath. 1996. “Migration with Endogenous Moving Costs.” *American Economic Review* 86(4): 909-30.

Chan, Kam Wing and Li Zhang. 1999. “The Hukou System and Rural Urban Migration in China Processes and Changes.” *China Quaterly* 160: 818-55.

Chen, Zhao, Shiqing Jiang, Ming Lu, and Hiroshi Sato. 2008. “How do Heterogeneous Social Interactions affect the Peer Effect in Rural–Urban Migration: Empirical Evidence from China.” LICOS Discussion Papers 22408, Centre for Institutions and Economic Performance, K.U. Leuven.

Cipollone，Piero, and Alfonso Rosolia. 2007. “Social Interactions in High School: Lessons from an Earthquake.” *American Economic Review* 97(3): 948-65.

Clampet-Lundquist, Susan and Douglas S. Massey. 2008. “Neighborhood Effects on Economic Self-Sufficiency: A Reconsideration of the Moving to Opportunity Experiment.” *American Journal of Sociology* 114(1):107-43.

Clark, Andrew E., and Youenn Lohéac. 2005. “‘It Wasn’t Me, It Was Them!’ Social Influence in Risky Behavior by Adolescents.” IZA Discussion Papers 1573.

Conley, Timothy G., and Christopher R. Udry. 2010. “Learning about a New Technology: Pineapple in Ghana.” *American Economic Review* 100(1):35–69.

Contreras, Dante, Diana Kruger, Marcelo Ochoa, and Daniela Zapata. 2007. “The Role of Social Networks in the Economic Opportunities of Bolivian Women.”Working paper.

Corcoran, Mary, Linda Datcher Loury, and Greg Duncan. 1980. “Information and Influence Networks in Labor Markets”, pp. 1-37 in G. J. Duncan and J. N. Morgan (eds) *Five Thousand American Families: Patterns of Economic Progress.* Ann Arbor, MI: Institute for Social Research.

Curran, Sara R., and Rivero-Fuentes E. 2003. “Engendering Migrant Networks: The Case of Mexican Migration”, *Demography* 40: 289-307.

Currie, Janetm, and Yelowitz, Aaron. 2000. “Are Public Housing Projects Good for Kids?” *Journal of Public Economics* 75(1): 99-124.

Cutler, David M, and Edward L. Glaeser. 1997. “Are Ghettos Good or Bad?” *The Quarterly Journal of Economics* 112(3):827-72.

De Graaf, Nan Dirk, and Henk D. Flap. 1988. “With a Little Help from My Friends: Social Resources as an Explanation of Occupational Status and Income in West Germany, the Netherlands, and the United States.” *Social Forces* 67:452-72.

Delattre, Eric, and Mareva Sabatier. 2007. “Social Capital and Wages: An Econometric Evaluation of Social Networking’s Effects.” *Labour* 21 (2): 209-36.

De Brauw, Alan, and Giles, John, 2008. “Migrant Labor Markets and The Welfare of Rural Households in The Developing World: Evidence From China.” Policy Research Working Paper Series 4585, The World Bank.

DiPrete, Thomas A., and Markus Gangl. 2004. “Assessing Bias in the Estimation of Causal Effects: Rosenbaum Bounds on Matching Estimates and Instrumental Variables with Imperfect Instruments.” *Sociological Methodology* 34(1): 271-310.

Domínguez, Silvia and Celeste Watkins. 2003. “Creating Networks for Survival and Mobility: Social Capital Among African-American and Latin-American Low-income Mothers.” *Social Problems* 50:111-135.

Dujardin, Claire, and Florence Goffette-Nagot. 2007. “Neighborhood Effects, Public Housing and Unemployment in France”. Working Papers 0505, Groupe d’Analyse et de Thèorie Economique (GATE), Universit´e Lyon 2.

Duflo, Esther, and Emmanuel Saez. 2003. “The Role Of Information And Social Interactions In Retirement Plan Decisions: Evidence From A Randomized Experiment.” *The Quarterly Journal of Economics* 118(3):815–42.

Duncan, Greg J., Jeanne Brooks-Gunn, and Pamela Klebanov. 1994. “Economic Deprivation and Early Childhood Development.” *Child Development* 65(2): 296-318.

Duncan, Greg J., James P. Connell, and Pamela K. Klebanov, 1997. “Conceptual and Methodological Issues in Estimating Causal Effects of Neighbourhoods And Family Conditions on Individual Development.”In J. Brooks-Gunn .G. Duncan. J. Aber (eds.) *Neighbourhood Poverty: Context and Consequences for Children 1*. New York: Russell Sage

Duncan, Greg J., Elizabeth Clark-Kauffman, and Emily Snell. 2004. “Residential Mobility Interventions as Treatments for the Sequelae of Neighborhood Violence.” Working paper,

Northwestern University.

Duncan, Otis D., Archibald O. Haller, and Alejandro Portes 1968. "Peer Influences on Aspirations: A Reinterpretation." *American Journal of Sociology* 74: 119-37.

Durlauf, Steven N. 1999. "The Case 'Against' Social Capital." *Focus* 20:1-4.

----2002. "On the Empirics of Social Capital". *Economic Journal*, 112(483):479-59.

Durlauf, Steven N. and Marcel Fafchamps. 2004. "Social Capital." NBER Working Paper 10485.

Durlauf, Steven N. and Hisatoshi Tanaka. 2008. "Understanding Regression Versus Variance Tests For Social Interactions." *Economic Inquiry* 46(1):25-28

Elliott, James R. 1999. "Social isolation and labor market insulation: Network and neighborhood effects of less-educated urban workers." *The Sociological Quarterly* 40:199–216.

Erickson, Bonnie H. 1996. "Culture, Class, and Connections". *American Journal of Sociology* 102(1): 217-51.

Evans, William N., Wallace E. Oates, and Robert M. Schwab. 1992. "Measuring Peer Effects：A Study of Teenage Behavior." *Journal of Political Economics* 100:966–91.

Falcon, Luis, and Edwin Melendez. 2001. "The Social Context of Job Searching for Racial Groups in Urban Centers." In: Alice O'Connor, Chris Tilly and Lawrence D. Bobo (eds.), New York. Sage: *Urban Inequality: Evidence from Four Cities*.

Falk, Armin, and Urs Fischbacher. 2002. "'Crime' in the Lab-detecting Social Interaction." *European Economic Review* 46(4-5): 859-69.

Fawcett, James T. 1989. "Networks, Linkages, and Migration Systems." *International Migration Review* 23: 671-80.

Feng, Shuaizhang, Alan B. Krueger, and Michael Oppenheimer. 2010. "Linkages Among Climate Change, Crop Yields and Mexico–US Cross-border Migration." *Proceedings of the National Academy of Science.* 107(32):14257-14262.

Flap, Henk, and Beate Völker. 2001, "Goal Specific Social Capital and Job Satisfaction." *Social Networks* 23:297–320.

Flap, Henk, and Ed Boxman. 2001. "Getting Started: the Influence of Social Capital on the Start of the Occupational Career". Pp. 159–81 in N. Lin, K. Cook and R.S. Burt (ed.) *Social*

Capital: Theory and Research. New York: Aldine De Gruyter.

Freese, Jeremy. 2007a. "Replication Standards For Quantitative Social Science: Why Not Sociology?" *Sociological Methods and Research* 36(2):153-72.

------ 2007b. "Overcoming Objections to Open-Source Social Science." *Sociological Methods and Research* 36(2): 220–26.

Galbiati，Roberto, and Giulio Zanella. 2008. "The Social Multiplier of Tax Evasion: Evidence from Italian Audit Data." Department of Economics University of Siena 539.

Gerber, Theodore P. 2000. "Membership Benefits or Selection Effects? Why Former Communist Party Members Do Better in Post-Soviet Russia." *Social Science Research* 29:25–50.

Glaeser, Edward L., Bruce Sacerdote, and Jose A. Scheinkman. 1996. "Crime and Social Interactions." *The Quarterly Journal of Economics* 111(2): 507-48.

Gong, Qianwen, and Junbiao Zhang. 2007. "Relations between Agricultural Natural Disasters and Rural Poverty Based on Panel Analysis in Anhui Province." *Chinese Population, Resources and Enviroment*, 17(4).

Gould, Eric D., Victor Lavy, and M. Daniele Paserman. 2005. "Does Immigration Affect the Long-Term Educational Outcomes of Natives? Quasi-Experimental Evidence." IZA Discussion Papers 1883

Granovetter, Mark S. 1974:1995. *Getting a Job: A Study of Contacts and Careers*. Cambridge, MA:Harvard University Press.

------1973. "The Strength of Weak Ties." *American Journal of Sociology* 78(6): 1360-80.

------1985. "Economic Action and Social Structure: the Problem of Embeddedness.", *American Journal of Sociology* 91(3):481-510.

Grodner, Andrew, and Thomas J. Kniesner 2007. "Labor Supply with Social Interactions: Econometric Estimates and Their Tax Policy Implications." IDEAS Working Paper.

Grogger, Jeff. 1996. "School Expenditures and Post-schooling Earnings：Evidence from the High School and Beyond." *Review of Economics and Statistics* 78(4):628-37.

Gurak, Douglas T., and Fe Caces. 1992. "Migration Networks and the Shaping of International Migration Systems." In Mary Kritz, Lin Lean Lim, and Hama Zlotmk, (eds.)

International Migration Systems: A Global Approach. Oxford: Clarendon Press: 150-76.

Harding, David J. 2003. "Counterfactual Models of Neighborhood Effects: The Effect of Neighborhood Poverty on Dropping Out and Teenage Pregnancy." *American Journal of Sociology* 109: 676-719.

Hare, Denise. 1999. "'Push' versus 'Pull' Factors in Migration Outflows and Returns: Determinants of Migration Status and Spell Duration among China's Rural Population." *Journal of Development Studies* 35(3): 45-72.

Halaby, Charles N. 2004. "Panel Models in Sociological Research: Theory into Practice." *Annual Review of Sociology* 30:507-44.

Hannan, Kate. 2008. "China: Migrant Workers Want 'Decent Work'." The Copenhagen Journal of Asian Studies, 26(2):60-81.

Hare, Denise, and Shukai Zhao. 2000. "Labor Migration as a Rural Development Strategy: A View from the Migration Origin." In West and Zhao (eds.) *Rural Labor Flows in China*, Institute of East Asian Studies, University of California, Berkeley.

Harding, David J. 2003. "Counterfactual Models of Neighborhood Effects: The Effect of Neighborhood Poverty on Dropping Out and Teenage Pregnancy." *American Journal of Sociology* 109:676–719.

Harris John R., and Michael P. Todaro. 1970. "Migration, Unemployment and Development: A Two-Sector Analysis." *American Economic Review* 60(1) 126-42.

Hartmann, Wesley R., Puneet Manchanda, Harikesh Nair, Matthew Bothner, Peter Dodds, David Godes, Kartik Hosanagar, and Catherine Tucker. 2008. "Modeling Social Interactions: Identification, Empirical Methods, and Policy Implications." *Marketing Letters* 19(3):287-304.

Hartmann, Wesley R., and Taylan Yildiz. 2006. "Preferences for Joint Consumption and the Implications for Targeted Marketing." Working paper.

Haug, Sonja 2008. "Migration Networks and Migration Decision-Making." *Journal of Ethnic and Migration Studies* 34(4):585-605.

Hanushek, Eric A., John F. Kain, Jacob M. Markman, and Steven G. Rivkin. 2003. "Does Peer Ability Affect Student Achievement?" *Journal of Appied Econometrica* 18(5): 527-44.

Hanushek, Eric A., Steven G. Rivkin, and Lori L. Taylor. 1996. "Aggregation and The

Estimated Effects of School Resources." *Review of Economics and Statistics* 78(4):611-27.

Heckman James. J. 1976. "The Common Structure of Statistical Models of Truncation, Sample Selection and Limited Dependent Variables and A Simple Estimator for Such Models." *Annals of Economic and Social Measurement* 5:475-92.

Heckman James. J. 1979. "Sample Selection Bias as a Specification Error." *Econometrica*. 47(1): 153-61.

Heckman James. J., and Richard Robb, Jr. 1985. Alternative Methods for Evaluating the Impact of Interventions: An Overview. *Journal of Econometrics*, 30(1-2) 239-67.

Hedström, Peter, and Richard Swedberg. 1998. *Social Mechanisms: An Analytical Approach to Social Theory.* Cambridge: Cambridge University Press.

Hedström, Peter. 2005. *Dissecting the Social: On the Principles of Analytical Sociology*. Cambridge: Cambridge University Press.

---- 2008. "Studying Mechanisms To Strengthen Causal Inferences In Quantitative Research." Forthcoming in J. M. Box-Steffensmeier, H. E. Brady & D. Collier (eds.) *The Oxford Handbook of Political Methodology*. Oxford:Oxford University Press.

Hedström, Peter, Yvonne Åberg, and Ann-Sofie Kolm. 2003. "Social Interactions and Unemployment". Working Paper.

Hellerstein, Judith K., Melissa McInerney, and David Neumark. 2009. "Neighbors And Co-workers: The Importance of Residential Labor Market Networks". NBER Working Paper 14201.

Holland, Paul W. 1986. "Statistics and Causal Inference." *The Journal of the American Statistical Association* 81 (396): 945-960.

Holzer, Harry J. 1987. "Hiring Procedures in the Firm: Their Economic Determinants and Outcomes." in *Human Resources and Firm Performance*. Madison, WI: Industrial Relations Research Association: 243-72.

Hoxby, Caroline M. 2000. "Peer Effects in the Classroom: Learning from Gender and Race Variation." NBER Working paper 7867.

Hoxby, Caroline M., and Gretchen Weingarth. 2005. "Taking Race Out of the Equation: School Reassignment and the Structure of Peer Effects." Working Paper. Harvard University.

Hughes, Marion R. 1997. "Sample Selection Bias in Analyses of the Political Democracy

and Income Inequality Relationship." *Social Forces* 75(3):1101-17.

Hunter, Lori M. 2005. "Migration and Environmental Hazards." *Population and Environment,* 26 (4): 273–302.

Hwang, Kwang-kuo.1987. "Face and Favor: The Chinese Power Game". *American Journal of Sociology* 92 (4) : 944-74.

Ichino, Andrea, and Giovanni Maggi. 2000. "Work Environment And Individual Background: Explaining Regional Shirking Differentials In A Large Italian Firm." *The Quarterly Journal of Economics* 115(3)**:**1057-90.

Imbens, Guido W., and Joshua D. Angrist. 1994. Identification and Estimation of Local Average Treatment Effects, *Econometrica* 62(2): 467-75.

Ioannides, Yannis M., and Linda Datcher Loury. 2004. "Job Information Networks, Neighborhood Effects, and Inequality." *Journal of Economic Literature* 42(4):1056-93.

Jacob, Brian A. 2004. "Public Housing, Housing Vouchers and Student Achievement: Evidence from Public Housing Demolitions in Chicago." *American Economic Review* 94(1):233-58.

Jencks, Christopher, and Susan E. Mayer. 1990. "The Social Consequences of Growing Up in a Poor Neighborhood." In Lynn, L.E., McGeary, M.G.H. (Eds.) *Inner-City Poverty in the United States.* Washington,DC: National Academy Press.

Kandel, Denise B.1978. "Homophily, Selection， and Socialization in Adolescent Friendships." *American Journal of Sociology* 84(2): 427-36.

Katz，Lawrence F., Jeffrey R. Kling, and Jeffrey B. Liebman. 2001. "A Moving to Opportunity in Boston: Early results of a randomized mobility experiment." *Quarterly Journal of Economics* 116: 607–54.

Kawaguchi, Daiji. 2004. "Peer Effects on Substance Use Among American Teenagers." *Journal of Population Economics* 17(2): 351–67.

Kim, Harris 2001. "Social Capital, Embedded Status, and the Endorsement Effect." Paper prepared for the SSRC workshop on the Corporation as a Social Institution, Institute of Industrial Relations, UC Berkeley.

Kling, Jeffrey R., Jeffrey B. Liebman, and Lawrence F. Katz. 2007. "Experimental Analysis of Neighborhood Effects." *Econometrica* 75(1):83-119.

Kling, Jeffrey R., Jeffrey B. Liebman, Lawrence F. Katz, and Lisa Sanbonmatsu. 2004. "Moving to Opportunity and Tranquility: Neighborhood Effects on Adult Economic Self-Sufficiency and Health from a Randomized Housing Voucher Experiment." Working Paper. Princeton University.

Kling, Jeffrey R., Jens Ludwig, and Lawrence F. Katz. 2004. "Neighborhood Effects on Crime for Female and Male Youth: Evidence from a Randomized Housing Mobility Experiment." *Quarterly Journal of Economics* 120:87-130.

Korenman, Sanders, and Susan C. Turner. 1996. "Employment Contacts and Minority-White Wage Differences." Industrial Relations 35: 106-22.

Korpi, Tomas. 2001. "Good Friends in Bad Times? Social Networks and Job Search among the Unemployed in Sweden." *Acta Sociologica* 44(2):157-70.

Krauth, Brian. 2004. "Peer Effects and Selection Effects on Smoking among Canadian Youth."*Canadian Journal of Economics* 38(3):735-57.

Kugler, Adriana D. 2003. "Employee Referrals and Efficiency Wages." *Labour Economics,* 10(5): 531–56.

Laschever, Ron 2005, "The Doughboys Network: Social Interactions and the Labor Market Outcomes of World War I Veterans." mimeo, Northwestern University.

Lavy, Victor, Daniel Paserman, and Analia Schlosser. 2008. "Inside the Black Box of Ability Peer Effect: Evidence from Variation of Low Achiever in the Classroom." NBER working paper No 14415.

Lazarsfeld, Paul F., and Robert K. Merton. 1954. "Friendship as a Social Process" Pp. 18-66 in M. Berger, T. Abel, and C. H. Page (ed.), *Freedom and Control in Modern Society*. Princeton, N.J.: Van Nostrand.

Laumann, Edward O. 1966. *Prestige and Association in an Urban Community: An Analysis of an Urban Stratification System*. Indianapolis: Bobbs.

Li, Bobai, and Andrew Walder. 2001. "Career Advancement as Party Patronage: Sponsored Mobility into the Chinese Administrative Elite, 1949-1996". *American Journal of Sociology* 106(5): 1371-1408.

Li, Chunlin. 2005. "Prestige Stratification in the Contemporary China: Occupational Prestige Measures and Socio-Economic Index." *Sociological Research* 2:74-102 (in Chinese).

Li, Shi. 2008. Rural Migrant Workers in China: Scenario, Challenges and Public Policy, Working Paper.

Liang, Zai. 2006. "Internal Migration in China in the Reform Era: Patterns, Policies, and Challenges." Pp. 197-215, In Zhongwei Zhao and Fei Guo (eds.) *Demography in China in the 21st Century.* Oxford University Press.

Liebman, Jeffrey B., Lawrence F. Katz, and Jeffrey R. Kling. 2004. "Beyond Treatment Effects: Estimating the Relationship between Neighborhood Poverty and Individual Outcomes in the MTO Experiment." Working Paper, Princeton University.

Lin, Nan. 1982. "Social Resources and Instrumental Action". Pp. 131–145, In P. V. Marsden and N. Lin (ed.) *Social Structure and Network Analysis.* Beverly Hills, CA: Sage

---- 1990. "Social Resources and Social Mobility: A Structural Theory of Status Attainment". Pp. 247–271. In R.L. Breiger (ed.) *Social Mobility and Social Structure*. New York: Cambridge University Press.

----1999. "Social Networks and Status Attainment". *Annual Review of Sociology.* 25:467-487.

----2001. "Building a Network Theory of Social Capital." In R. Burt, K. Cook, and N. Lin (eds.) *Social Capital: Theory and Research,* eds. New York: Aldine De Gruyter

----2003. "Job Search in Urban China: Gender, Network Chains, and Embedded Resources". Pp. 145-171. In Henk Flap and Beate Völker (ed.) *Creation and Return to Social Capital*. New York: Praeger.

----2008. "A Network Theory of Social Capital". Pp. 50-59. In Dario Castiglione, Jan van Deth, and Guglielmo Wolleb (ed.) *The Handbook of Social Capital.* Oxford and New York: Oxford University Press.

Lin, Nan, and Dan Ao. 2008. "The Invisible Hand of Social Capital: An Exploratory Study." In Lin and Ericson. 2008. *Social Capital: An International Research Program.* Oxford: Oxford University Press.

Lin, Nan, and Mary Dumin. 1986. "Access to Occupations Through Social Ties". *Social Networks* 8: 365–85.

Lin, Nan, and Wen Xie. 1988. "Occupational Prestige in Urban China." *American Journal of Sociology.* 93:793-832.

Lin, Nan, and Yanjie Bian. 1991. "Getting Ahead in Urban China". *American Journal of Sociology* 97:657-688.

Lin, Nan, Walter M. Ensel, and John C. Vaughn. 1981. "Social Resources and Strength of Ties: Structural Factors in Occupational Status Attainment." *American Sociological Review* 46 (4): 393-405.

Lin, Nan, John C. Vaughn, and Walter M. Enseld 1981. "Social Resources and Occupational Status Attainment . *Social Forces,* 59(4): 1163-81.

Louch, Hugh. 2000. "Personal Network Integration: Transitivity and Homophily in Strong-tie Relations". *Social Networks*, 22:45-64.

Loury, Linda Datcher. 2006. "Some Contacts Are More Equal than Others: Earnings and Job Information Networks." *Journal of Labor Economics* 24: 299-318.

Ludwig, Jens, Greg J. Duncan, and Joshua C. Pinkston. 2005. "Housing Mobility Programs and Economic Self-sufficiency: Evidence from a randomized experiment." *Journal of Public Economics* 89(1):131-56.

Mallee, Hein. 2000. "Agricultural Labor and Rural Population Mobility: Some Observations." In West and Zhao (eds.) *Rural Labor Flows in China*, Institute of East Asian Studies, University of California, Berkeley.

Manski, Charles F. 1993. "Identification of Endogenous Social Effects: The Reflection Problem." *Review of Economic Studies* 60(3):531-42.

------ 2000. "Economic Analysis of Social Interactions." *Journal of Economic Perspectives* 14:115-136.

Margolis, David N., and Véronique Simonnet. 2003. "Educational Track, Networks and Labor Market Outcomes." IZA Discussion Paper No. 699.

Marmaros, David, and Bruce Sacerdote. 2002 "Peer Effects in Occupational Choice." *European Economic Review* 46: 870-79.

Marsden, Peter V. 2001. "Interpersonal Ties, Social Capital, and Employer Staffing Practices." Pp:105-25 in N. Lin, K. S. Cook, and R. Burt (eds.) *Social Capital: Theory and Research.* Hawthorne, NY: Aldine de Gruyter.

Marsden, Peter V., and Elizabeth Gorman. 2001. "Social Networks, Job Changes, and Recruitment" Pp: 467–502. In Berg and Kalleberg (eds), *Sourcebook on Labor Markets:*

Evolving Structures and Processes. New York: Kluwer Academic/Plenum Publishers.

Marsden, Peter V., and Jeanne Hurlbert. 1988. "Social Resources and Mobility Outcomes: A Replication and Extension." *Social Forces* 66(4): 1038-1059.

Massey, Douglas S. 1987. "Understanding Mexican Migration to the United States." *American Journal of Sociology* 92(6), 1372-1403.

Massey, Douglas S., Joaquin Arango, Graeme Hugo, Ali Kouaouci, Adela Pellegrino, and J. Edward Taylor. 1993. "Theories of International Migration: A Review and Appraisal." *Population and Development Review* 19(3):431-66.

Massey, Douglas S., and Kristin E. Espinosa. 1997. "What's Driving Mexico-U.S. Migration? A Theoretical, Empirical, and Policy Analysis." American Journal of Sociology 102(4) 939–99.

Maurin, Eric, and Julie Moschion. 2009. "The Social Multiplier and Labor Market Participation of Mothers." *American Economic Journal: Applied Economics* 1(1):251-72.

McEwan, Patrick J. 2003. "Peer Effects on Student Achievement: Evidence from Chile." *Economics of Education Review* 22(2):131-41.

McLeman, Robert A., and Barry Smit. 2006. "Migration as an Adaptation to Climate Change." *Climatic Change* 76 (1-2): 31-53.

McPherson, J. Miller, Lynn Smith-Lovin, and James M. Cook. 2001. "Birds of a Feather: Homophily in Social Networks". *Annual Review of Sociology* 27:415–44.

Meng, Xin. 2000. *Labour Market Reform in China.* Cambridge: Cambridge University Press.

Meyerson, Eva. M. 1994. "Human Capital, Social Capital and Compensation: The Relative Contribution of Social Contacts to Managers' Incomes." *Acta Sociologica* 37: 383-99.

Moffit, Robert A. 1995. "Selection Bias Adjustment in Treatment-effects Models as a Method of Aggregation." Pp: 234-238. In "1996 Proceedings of the American Statistical Association".

------2001. "Policy Interventions, Low-Level Equilibria and Social Interactions". In S. N. Durlauf, and P. H. Young (ed.) *Social Dynamic*. Boston, MA：MIT Press.

Mollenhorst, Gerald, Beate Völker, and Henk Flap. 2008. "Social Contexts and Core Discussion Networks: Using a Choice-Constraint Approach to Study Similarity in Intimate Personal Relationships." *Social Forces*, 86:937-965.

Montgomery, James D. 1991. “Social Networks and Labor-Market Outcomes: Toward an Economic Analysis.” *American Economic Review* 81 (Dec.): 1408-18.

----1992. Job Search and Network Composition: Implications of the Strength-of-Weak-Ties Hypothesis. *American Sociological Review* 57:586-596.

Morgan, Stephen L. 2001. “Counterfactuals, Causal Effect Heterogeneity, and the Catholic School Effect on Learning.” *Sociology of Education* 74(4):341-74.

------ 2002. “Should Sociologists Use Instrumental Variables?” Working Paper Cornell University.

Morgan, Stephen L., and Christopher Winship. 2007. Counterfactuals and Causal Inference: Methods and Principles for Social Research. Cambridge: Cambridge University Press.

Morgan, Stephen L., and David J. Harding. 2006. “Matching Estimators of Causal Effects: Prospects and Pitfalls in Theory and Practice.” *Sociological Methods and Research* 35:3-60.

Mosca, Michele, and Francesco Pastore. 2008. “Wage Effects of Recruitment Methods: The Case of the Italian Social Service Sector.” IZA Discussion Paper 3422.

Mouw, Ted. 2002. “Racial Differences in the Effects of Job Contacts: Conflicting Evidence from Cross-Sectional and Longitudinal Data.” *Social Science Research* 31(4): 511-38.

------ 2003. “Social Capital and Finding a Job: Do Contacts Matter?” *American Sociological Review* 68 (6): 868-98.

------ 2006. “Estimating the Causal Effect of Social Capital: A Review of Recent Research”. *Annual Review of Sociology* 32:53-78.

Munshi, Kaivan. 2003. “Networks in the Modern Economy: Mexican Migrants in the U.S. Labor Market.” *Quarterly Journal of Economics* 118(2):549-99.

Munshi, Kaivan, and Jacques Myaux. 2006. “Social Norms and the Fertility Transition.” *Journal of Development Economics* 80(1):1-38.

Nair, Harikesh, Puneet Manchanda, and Tulikaa Bhatia. 2006. “Asymmetric Peer Effects in Physician Prescription Behavior: The role of opinion leaders.” SSRN working paper.

Nee, Victor. 1989. “A Theory of Market Transition: From Redistribution to Markets in State Socialism.” American Sociological Review 54(5):663–81.

------ 1991. “Social Inequalities in Reforming State Socialism: Between Redistribution

and Markets in State Socialism." *American Sociological Review* 56(3):267–82.

------ 1996. "The Emergence of a Market Society: Changing Mechanism of Stratification in China." *American Journal of Sociology* 101(4):908–49.

Newcomb, Theodore M. 1961. *The Acquaintance Process.* New York: Holt, Rinehart and Winston.

Oreopoulous, Philip. 2002. "Do Neighbourhoods Influence Long-term Labour Market Success? A Comparison of Adults Who Grew Up in Different Public Housing Projects". *Statistics Canada, Analytical Studies Branch Research* 11F0019. No.:185.

Palloni, Alberto, Douglas S. Massey, Miguel Ceballos, Kristin Espinosa, and Michael Spittel. 2001. "Social Capital and International Migration：A Test Using Information on Family Networks." *American Journal of Sociology* 106 (5): 1262-98.

Petersen, Trond, Ishak Saporta, and Marc-David L. Seidel. 2000. "Offering a Job: Meritocracy and Social Networks." *American Journal of Sociology* 106(3): 763-816.

Plotnick, Robert, and Saul Hoffman. 1995. "Fixed Effects Estimates of Neighborhood Effects. Working paper. University of Washington

Podolny, J. M., & Baron, J. N. 1997. "Resources and Relationships: Social Networks and Mobility in the Workplace. *American Sociological Review,* 62 (5):673-693.

Ponzo, Michela, and Vincenzo Scoppa. 2009. 'The Use of Informal Networks in Italy: Efficiency or Favoritism?' MPRA Paper 17281, University Library of Munich.

Portes, Alejandro. 1998. "Social Capital: Its Origins and Applications in Modern Sociology." *Annual Review of Sociology* 24 (1): 1-24.

Qi, Xin. 2007. "Continuing and Rebuilding: the Social Network of Cainozoic Rural Migrant Workers." *The Journal Of Jiangsu Administration Institute* (3)27-38.(In Chinese)

Rainer, Helmut, and Tom Siedler. 2008. "Social Networks in Determining Migration and Labour Market Outcomes: Evidence from the German Reunification." ISER Working Papers 2008-36.

Renzuli, Linda A., and Howard Aldrich. 2005. "Who Can You Turn To? Tie Activation Within Core Business Discussion Networks". *Social Forces* 84 (1): 323-41.

Requena, Felix 1991. "Social Resources And Occupational Status Attainment In Spain: A

Cross-Ntional Comparison With The United States And The Netherlands." *International Journal of Comparative Sociology* 32(3-4):233-42.

Ritchey, P. Neal. 1976. "Explanations of Migration." *Annual Review of Sociology* 2: 363-404.

Rivkin, Steven. 2001. "Tiebout Sorting, Aggregation and the Estimation of Peer Group Effects." *Economics of Education Review* 20(3) 201-09.

Roberts, Kenneth D. 2000. "Chinese Labor Migration: Insights from Mexican Undocumented Migration to the United States," in West and Zhao (eds.) Rural Labor Flows in China, Institute of East Asian Studies, University of California, Berkeley.

Robins, Garry, and Jennifer Boldero.2003. "Relational Discrepancy Theory: The Implications of Self-discrepancy Theory for Dyadic Relationships and for the Emergence of Social Structure". *Personality and Social Psychology Review* 7(1): 56–74.

Robinson, Chris, and Nigel Tomes. 1982. "Self-selection and Interprovincial Migration in Canada." *Canadian Journal of Economics* 15: 474-502.

Rona-Tas, Akos, and Alya Guseva. 2001. "The Privileges of Past Communist Party Membership in Russia and Endogenous Switching Regression." *Social Science Research* 30:641–52.

Rosenbaum，Paul R., and Donald B. Rubin. 1983. "The Central Role of the Propensity Score in Observational Studies for Causal Effects." *Biometrika* 70(1):41-55.

Rozelle, Scott, Li Guo, Minggao Shen, John Giles and Tuan Yee Low. 1997. "Poverty, Networks, Institution, or Education: Testing among Competing Hypotheses on the Determinants of Migration in China." Paper presented at the 1997 Annual Meetings of the Association for Asian Studies, Chicago, Illinois, March 13-16.

Ruiter, Stijn, and Nan Dirk De Graaf. 2009. "Socio-economic Payoffs of Voluntary Association Involvement: A Dutch Life Course Study." *European Sociological Review*. 25: 425-442.

Sacerdote, Bruce. 2001. "Peer Effects With Random Assignment: Results with Dartmouth roommates." *The Quarterly Journal of Economics* 116(2) 681-704.

Sampson, Robert J. 2008. "Moving to Inequality: Neighborhood Effects and Experiments Meet Social Structure." *American Journal of Sociology* 114 (1):189-231.

Sanbonmatsu, Lisa, Jeffrey R. Kling, Greg J. Duncan, Jeanne Brooks-Gunn. 2006. "Neighborhoods and Academic Achievement: Results From the Moving to Opportunity Experiment." NBER Working Paper 11909.

Seibert, Scott E., Maria L. Kraimer, and Robert C. Liden. 2001. "A Social Capital Theory of Career Success". *The Academy of Management Journal* 44(2):291-237.

Shu, Xiaoling, and Yanjie Bian. 2003. "Market Transition and Gender Gap in Earnings in Urban China." *Social Forces* 81(4): 1107-45.

Shvydko, Tetyana 2007, "Interactions at the Workplace: Peer Erects in Earnings." Job Market Paper, Department of Economics and Carolina Population Center, University of North Carolina at Chapel Hill.

Simon, Curtis, and John Warner. 1992. "Matchmaker, Matchmaker: The Effect of Old Boy Networks on Job Match Quality, Earnings, and Tenure." *Journal of Labor Ecomics* 10(3): 306–30.

Smith, Sandra S. 2000. "Mobilizing Socialre Sources: Race, Ethnic, and Gender Differences in Social Capital and Persisting Wage Inequalities." *The Sociological Quarterly* 41(4):509-38.

Sobel, Michael E. 1996. "An Introduction to Causal Inference". *Sociological Methods and Research* 24: 353-79.

Soetevent, Adriaan R. 2006. "Empirics of the Identification of Social Interactions: An Evaluation of the Approaches and Their Results." *Journal of Economic Surveys* 20(2):193-228.

Soetevent, Adriaan R., and Peter Kooreman. 2005. "Social Ties within School Classes: The Roles of Gender, Ethnicity, and Having Older Siblings." Oxford Review of Economic Policy 21(3):373-91.

Solinger, Dorothy J. 1999. "Citizenship Issues in China's Internal Migration: Comparisons with Germany and Japan." *Political Science Quarterly* 114(3):455-78.

Sorensen, Alan T. 2006. "Social Learning and Health Plan Choice." *The RAND Journal of Economics.* 37(4): 929–45.

Sorensen, Jesper B. 2000. "The Longitudinal Effects of Group Tenure Composition on Turnover." *American Sociological Review* 65(2):298-310

Sprengers, Maarten, Fritz Tazelaar, and Hendrik Derk Flap. 1988. "Social Resources,

Situational Constraints, and Reemployment". Netherlands Journal of Sociology, 24:98–116.

Stock, James H., and Motohiro Yogo. 2005. "Testing for Weak Instruments in Linear IV Regression". In Stock, J.H., Andrews, D.W.K. (eds.), Identification and Inference for Econometric Models: A Festschrift in Honor of Thomas Rothenberg. Cambridge University Press, Cambridge: 80–108. (Ch. 5).

Stock, James H., and Mark W. Watson. 1997. *Introduction to Econometrics* (2nd edition). Addison Wesley.

Stolsenberg, Ross M., and Daniel A. Relles. 1997. "Tools for Intuition about Sample Selection Bias and its Correction." *American Sociological Review* 62(3):494-507.

Taylor, Edward J. 1986. "Differential Migration, Networks, Information and Risk." In Oe Stark (eds.) *Migration Ttheory, Human Capital and Development*. Greenwich, Conn.: JAI Press: 147-171.

Todaro, Michael P. 1969. "A Model of Labour Migration and Urban Employment in Less Developed Countries." *American Economic Review* 59, 138-48.

Topa, Giorgio. 2001. "Social Interactions, Local Spillovers and Unemployment." *Review of Economics Studies* 68(2): 261-95.

Uhlenberg, Peter. 1973. "Noneconomic Determinants of Nonmigration: Sociological Consideration for Migration Theory. " *Rural Sociology* 38: 296-311.

Van Der Gaag, M.P.J, and Thomas. A.B Snijders. 2004. Proposals for the Measurement of Individual Social Capital. In: Flap, H. and Völker B. (Eds.), Creation and returns of Social Capital.London , Routledge, pp.199-218.

Van der Klaauw, Bas, and Jan C. van Ours. 2003. "From Welfare to Work: Does the neighborhood matter?" *Journal of Public Economics* 87:957-85.

Vigdor, Jacob, and Thomas Nechyba. 2004. "Peer Effects in Elementary School: Learning from 'Apparent' Random Assignment." Working paper.

Vecchio, Robert. P. 1995. "The Impact of Referral Sources on Employee Attitudes: Evidence from a National Sample." *Journal of Management* 21(5): 953-965.

Verbrugge, Lois M. 1977. "The Structure of Adult Friendship Choices." *Social Forces* 56:576–97.

Völker, Beatie. 1995. "Changes in Personal Networks in the Former GDR." In *Should Auld*

Acquaintance Be Forgot...? The Institutions of Communism, the Transition to Capitalism and Personal Networks: the Case of East Germany. Amsterdam: Thesis Publishers

Völker, Beate, and Henk Flap. 1999. "Getting Ahead in the GDR: Social Capital and Status Attainment under Communism." *Acta Sociologica* 42(1): 17-34.

Walder, Andrew G. 1986, *Communist Neo-traditionalism: Work and Authority in Chinese Industry.* California: California University Press.

------ 2003. "Elite Opportunity in Transitional Economics". *American Sociological Review* 68(6): 899-916.

Walder, Andrew G., Bobai Li, and Donald J. Treiman. 2000. "Politics and Life Chances in a State Socialist Regime: Dual Career Paths into the Urban Chinese Elite, 1949 to 1996." *American Sociological Review* 65(2):191–209.

Wang, Guomin and Ye Zheng. 2007. *The Risk Management System of Natural Disaster in Agriculture in China,* Xinan Caijing Press (In Chinese).

Wei, Yiming, Ying Fan, Cong Lu, and Hsien-Tang Tsai 2004. "The Assessment of Vulnerability to Natural Disasters in China by Using the DEA Method." *Environmental Impact Assessment Review* 24: 427–39.

Weinberg, Bruce A., Patricia B. Reagan, and Jeffrey J. Yankow. 2004. "Do Neighborhoods Affect Hours Worked? Evidence from Longitudinal Data". *Journal of Labor Economics* 22(4):891–924.

Wegener, Bernd. 1991. "Job Mobility and Social Ties: Social Resources, Prior Job, and Status Attainment." *American Sociological Review* 56 (1): 60-71.

Wellman, Barry, and Milena Gulia. 1999. "The Network Basis of Social Support: A Network Is More Than the Sum of Its Ties". Pp. 83-118. In Barry Wellman (ed.) *Networks in the Global Village: Life in Contemporary Communities*. Boulder, CO: Westview.

Wielgosz, John B., and Susan Carpenter. 1987. "The Effectiveness of Alternative Methods of Searching for Jobs and Finding Them." *American Journal of Economics and Sociology* 46: 151-64.

Windizio, Michael. 2008. "Social Structures and Actors: the Application of Multilevel Analysis in Migration Research." *Romanian Journal of Population Studies* 1:113-38.

Winship, Christopher and Stephen Morgam. 1999. "The Estimation of Causal Effects

from Observational Data". *Annual Review of Sociology* 25: 659-706.

Winship, Christopher, and Robert D. Mare. 1992. "Models for Sample Selection Bias." *Annual Review of Sociology* 18:327-350.

Winship, Christopher and Michael Sobel. 2004. "Causal Inference in Sociological Studies." In M Hardy, A Bryman (ed.) *Handbook of Data Analysis*. Thousand Oaks.CA: Sage.

Wong, Raymond Sin-Kwok. 1996. "The Social Composition of the Czechoslovak and Hungarian Communist Parties in the 1980s." *Social Forces* 75:61–90.

Wooldridge, Jeffrey M. 2002. *Econometric Analysis of Cross Section and Panel Data.* Cambridge，MA：MIT Press.

------ 2006. *Introductory Econometrics: A Modern Approach.* (Third Edition). Thomson South-Western.

Wu, Xiaogang, and Donald J. Treiman. 2007. "Inequality and equality under Chinese Socialism: the Hukou System and Intergenerational Occupational Mobility." *American Journal of Sociology* 113 (2): 415–45.

Wu, Xiaogang, and Yu Xie. 2003. "Does the Market Pay Off? Earnings Returns to Education in Urban China." *American Sociological Review* 68(3): 425-42.

Xu, Bingkui. 2007. "Social Network of Peasant Migrants in City and Their Interpersonal Communication." *The Journal of East China University of Science and Technology 3*.(In Chinese)

Xu, Xinxin. 2000. "Changes in the Chinese Social Structure as Seen from Occupational Prestige Ratings and Job Preferences" . Sociological Studies 3:67-85 (in Chinese).

Yakubovich, Valery. 2005. "Weak Ties, Information, and Influence: How Workers Find Jobs in a Local Russian Labor Market". *American Sociological Review* 70: 408–21.

Zhang, Li. 2001. *Strangers in the City: Reconfiguration of Space, Power and Social Networks within China's Floating Population.* Stanford: Stanford University Press.

Zhao, Yaohui. 1999. "Leaving the Countryside: Rural to Urban Migration Decisions in China." *American Economic Review* 89: 281-86.

------ 2002. "Causes and Consequences of Return Migration: Recent Evidence from China." *Journal of Comparative Economics* 30: 376-96.

------ 2003. "The Role of Migrant Networks in Labor Migration: The Case of China."

Contemporary Economic Policy 21(4): 500-11.

Zhu, Nong. 2002. "The Impact of Income Gaps on Migration Decisions in China." *China Economic Review* 13: 213-30.

Zhou, Xueguang. 2000. "Economic Transformation and Income Inequality in Urban China: Evidence from Panel Data." *American Journal of Sociology*. 105(4): 1135-74.

Zhou, Xueguang, and Olga Suhomlinova. 2000. "Redistribution under State Socialism: A USSR and PRC Comparison." *Research in Social Stratification and Mobility* 18: 163-204.

Zimmerman**,** David J. 2003. "Peer Effects in Academic Outcomes: Evidence from a natural experiment." *Review of Economics and Statistics* 85(1): 9-23.