APPLIED ECONOMETRIC STUDIES ON NETWORK EFFECTS

A Dissertation

by

SHIXUAN LI

Submitted to the Office of Graduate and Professional Studies of
Texas A&M University
in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

Chair of Committee,	Yonghong An
Committee Members,	Zheng Fang
	Qi Li
	Steven Puller
	Ximing Wu
Head of Department,	Timothy J. Gronberg

May 2020

Major Subject: Economics

ABSTRACT

This dissertation contains three essays which examine network effect study in econometrics and its applications. In the first essay, I study two research questions: how underwriter network affects firm network through IPO process and how firm network affects issuer post-IPO stock market performance. To address the two research questions, I construct firm and underwriter network based on sharing the same institutional shareholders and cooperating in the same syndicate, respectively. The networks are measured by network centrality measures from social network analysis (SNA): degree, eigenvector, closeness and betweenness. I then adopt regression models with these centrality measures. Empirical results indicate that an issuer is significantly more central in its public firm network if this issuer is led by a bookrunner through IPO process that is more central in its underwriter network. The effect of firm network on issuer post-IPO market performance is significant. An issuer with higher network centrality will achieve higher holding period return and higher monthly average trading volume in the first post-IPO year, while an issuer with higher eigenvector centrality has lower holding period return in the third post-IPO year. To analyze the source of underwriter network effect on firm performance, I run a specification by adding underwriter network centralities along with firm network centralities to analyze the firm post-IPO market performance. I also find that the effect of underwriter network on firm performance that is documented in literature can be mainly attributed to the effect of underwriter network on firm network.

In the second essay, we study identification and estimation of peer effects in observed networks where the problem of mismeasured links arises. In applied work, researchers generally construct networks that contain mismeasured links because of the data they use or the method of construction they adopt. Failing to deal with these mismeasured links, that using the observed networks as the true networks of interest, can result in biased estimates of peer effects parameters. Our paper provides sufficient conditions to identify peer effects in a generalized linear-in-mean model in which networks of interest contain mismeasured links. With the help of repeated observations of networks, we can identify peer effects in a generalized method-of-moments(GMM) model through

an observed conditional moment model, instead of knowing the latent network structures. Based on the identification strategy, we propose a three-stage semiparametric estimator and apply our method to analyze peer effects among firms on their financial policies. Empirical results indicate that the peer effect is significantly positive and it becomes insignificant when mismeasured links are not addressed. Empirically studying firm networks with mismeasured links provides us an in-depth look on the biased estimates of peer effects if mismeasured links are not properly taken care of.

In the third essay, we conduct a reduced-form analysis of first-price sealed-bid auction with affiliated private values under a network perspective. In recent literature, researchers generally adopt symmetric dependency among bidder's private values through affiliation. Failing to take the embedded network structure of those bidders into consideration, that assuming the dependent structures of every pair of bidders be the same, can result in biased estimates in dependency parameters. Our paper provides some empirical evidence on the dependency of private values among bidders based on their linked status and proposes a structural model for estimation. The primary data we use are collected from the detailed bid summary files provided by the California Department of Transportation (Caltrans). We construct contractor networks based on if two contractors share a same subcontractor in one project. In the end, reduced-form results indicate that being linked has a significantly negative effect on the difference of bid amount between two contractors (bidders). In other words, linked contractors tend to submit close bid prices.

DEDICATION

To my parents and my husband.

ACKNOWLEDGEMENTS

This dissertation is thanks in large part to the special people who challenged, supported and stuck with me along the way. I am tremendously fortunate to have committee members Prof.Yonghong An, Prof.Zheng Fang, Prof.Qi Li, Prof.Steven Puller and Prof.Ximing Wu. I thank them for supporting this project and giving such thoughtful feedback.

I want to express my sincere thanks to my committee chair Yonghong for bringing a depth of knowledge and giving me the most support in my research. I am extremely grateful and indebted to him for his expert, sincere and valuable guidance.

I also want to thank my coauthors, Dr.Wenzheng Gao and Li Zheng, for their patience in many discussions, and their supports, suggestions and opinions.

Last of all, I would like to thank my family and friends who supported me in all conditions. My parents love me and always stand by my side. My husband Dr.Guoyu Fu helped me a lot by sharing his knowledge and experience gained from his PhD life. My best friend Dr.Lu Yang, gave me many comforts and encouragements whenever I run into sadness.

The completion of this project could not have been possible without these people and many others that ever helped me whose names may not all be enumerated. Thank you all!

CONTRIBUTORS AND FUNDING SOURCES

Contributors

This work was supported by a dissertation committee consisting of Professor Yonghong An (chair), Professor Zheng Fang, Professor Qi Li, and Professor Steven Puller of the Department of Economics, and Professor Ximing Wu of the Department of Agricultural Economics.

The analyses depicted in Section Three were conducted in part by Dr.Wenzheng Gao. The analyses depicted in Section Four were conducted in part by Li Zheng.

All other work conducted for the dissertation was completed by the student independently.

Funding Sources

Graduate study was supported by the Department of Economics of Texas A&M University.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

1 INTRODUCTION

Network effect has been widely studied. Researchers are interested in how network structure may affect network outcomes. Network structure and network content are important factors for various aspects of network outcomes (e.g. see Burt (2000), Cohen and Frazzini (2008)). In this dissertation, we studied econometric methods under the topic of network effects and apply our method to different empirical works. There are three essays in this dissertation.

In the first essay, the objective of this paper is to analyze how the central position of a bookrunner in its underwriter network affects the position of an issuer in its public firm network in IPO process, and to study how the established position of the issuer in public firm network influences its post-IPO market performance. In the existing literature, effect has been seen of underwriter network on firm IPO performance (Bajo et al. (2016)). Nevertheless, the mechanism of underwriter network on firm performance is unclear. To further investigate the mechanism behind the effect, I study two research questions. The first question I analyze is how underwriter network affects firm network through IPO process. During book running process, bookrunner can induce institutional holders to pay attention to the firms it takes public by delivering noisy information to and collecting confidential bids from institutional holders. Intuitively, a bookrunner with higher network centrality in underwriter network has more sources on institutional holders, and can thus attract more investor attention, leading to higher network centrality of the issuer it takes public. Second, I examine how firm network affects issuer post-IPO market performance. Such effect could exist because institutional holders can affect firm's management decision and financial policy by setting board members. Moreover, institutional holders can deliver noisy information related to the firm to other firms through actions they take, such as selling a firm's shares or buying new shares. To empirically address the two questions, I first construct firm and underwriter network, then use reduced form analysis. To construct the two networks, I use data from Thomson Reuters 13f institutional holdings and Thomson Reuters SDC New Issues Database. To capture firm performance, IPO characteristics, and firm characteristics, data from Center for Research in Security Prices (CRSP) and Compustat are used. Firm network is constructed by defining the linkages between two firms based on if they share a same institutional holder. Following the literature, I define underwriter network in the way that two underwriters are linked if they ever cooperated in a same syndicate in the previous five years before IPO. My dependent variables are three post-IPO stock market performance measures: average monthly trading volume, holding period return in the first post-IPO year and holding period return in the third post-IPO year. My controls are variables that capture firm characteristics, IPO characteristics, and market environments. I adopt two regression models of network centrality measures as independent variables. The centrality measures are degree, eigenvector, closeness, and betweenness, from social network analysis (SNA). The first model is a linear regression of firm network centralities on underwriter networks centralities. The second model is a linear regression of firm post-IPO performance on firm network centralities.

In the second essay, the objective of this paper is to both theoretically and empirically study the identification and estimation strategy of peer effects in networks, given that the econometricians may not be able to observe the latent network because of mismeasured links. To identify peer effects in networks with mismeasured links, we take three steps. In the first step, we investigate the identification strategy in a generalized linear-in-mean social interaction model. We use two adjacency matrices that summarize the link structures in peer-effects network and contextual-effects network, separately. The outcome variable is linearly determined by the mean of neighbors' characteristics that in the contextual-effects network and the mean of peer outcomes that in the peer-effects network. Because of the linear-in-mean model, the misclassification of linkage status in the adjacency matrices are transformed to measurement errors in the mean outcomes and the mean characteristics. We first show that an weighted average of characteristics of neighbors in an individual's peers' contextual-effects network can be used as an instrument variable for the mean of that individual's peer outcomes (Bramoullé, Djebbari and Fortin (2009), Blume et al. (2015)). For simplicity, we call the instrument variable for the mean of peer outcomes as instrumental mean characteristics. The social effects can be identified if the observed network is assumed to be the true network of interest. Then to identify peer effects in networks with mismeasured links, we first apply the method in Hu (2008). We use repeated measurements on two conditional distribution matrices: one is the joint distribution matrix of observed mean characteristics/instrumental mean characteristics and outcomes conditional on covariates; the other is the distribution matrix of observed mean characteristics/instrumental mean characteristics conditional on covariates. Then we use an eigenvalue-eigenvector decomposition method based on above distribution matrices to obtain conditional densities of latent mean characteristics/instrumental mean characteristics conditional on observed mean characteristics/instrumental mean characteristics, separately.We use a GMM method to estimate peer effects parameters. In the second part, we develop a three-stage semiparametric estimator following the identification procedure. In the third part, we carry out an empirical study by constructing firm networks then using a structural form analysis. In our empirical study, to construct the peer-effects network and contextual-effects network, we use data from the Customer-supplier data and the merged Center for Research in Security Prices (CRSP)-Compustat database.

In the third essay, the objective of this paper is to empirically study the first-price sealed-bid auction with affiliation, given the network of potential bidders. To investigate the the first-price sealed-bid auction with affiliation under a network perspective, we take several steps. In the first step, we propose a simple model that characterize our basic settings and assumptions. For the isolated bidders, their private values are independent of others'. For the linked pairs of bidders, their private values are dependent. We further assume that the dependence structure of private valuations between any pair of linked bidders is the same, regardless of the auction and bidder characteristics. Next, we construct contract-specific contractor networks using data from California Department of Transportation (Caltrans). The contract-specific contractor network is constructed by the chance of having at least one same subcontractors in the underlying contract. Then based on the 712 contract-specific contractor networks from our sample, we investigate the densities of those networks. Moreover , we provide key evidence on how private values are dependent based on network structure. We first introduce the distance between two contractors and we find most pairs of contractors are either unlinked or within two-step reach. Then we construct a variable called Bid Amt, a ratio between total bid amount and engineer estimates. We investigate how the difference in Bid Amt between two contractors are determined by the distance between the two contractors. Lastly, we examine how the position of a contractor in its contractor network may affect its bidding behaviour and its probability of winning. We first run a linear regression of Bid Amt on four centrality measures (from social network analysis: degree, eigenvector, betweenness, closeness). Then we conduct a Probit model of probability of winning on centrality measures.

2 NETWORK CENTRALITY AND NETWORK EFFECTS

Network structure and network content are important factors for various aspects of network outcomes (e.g. see Burt (2000), Cohen and Frazzini (2008)). A network of entrepreneurial firms may affect accounting-based performance, market-based performance for public firms, or innovation and technology change, and so forth (e.g. see Deeds and Hill (1996), Cohen and Frazzini (2008), Ahern (2013)). Moreover, the network of entrepreneurial firms can be affected by other network, such as the underwriter network. Previous literature on inter-firm network focus on firm linkages as mergers, acquisitions, R&D alliances, customers and suppliers (e.g. Bengtsson and Kock (2000), Cohen and Frazzini (2008), Ahern and Harford (2014)). This paper studies the effect of underwriter network on firm network and the effect of firm network on firm post-IPO performance.

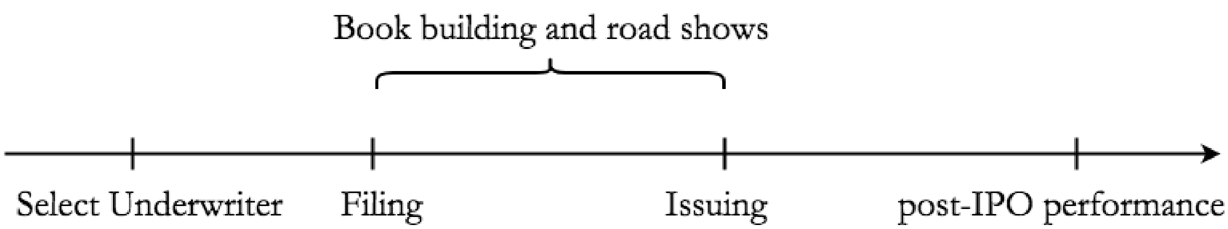


Figure 1: Timing of Initial Public Offering

In an initial public offering (IPO), two main parties involved are issuer (the firm that goes public) and bookrunner or lead underwriter (the investment bank that leads the IPO process and provides underwriting services for the issuer), as shown in Figure 1. Firms form linkages if they share the same large institutional holders. Intuitively, effectiveness of such linkages exists because of the two potential roles played by institutional holders: decision makers that can affect firm's management decisions or financial policies by setting board members in the firm, and paths along which the information of firms diffuse. Therefore, firm performance is potentially affected by its position in firm network. Underwriters form linkages if they cooperate in the same syndicate in the previous five years before IPO. Such linkages are effective because underwriters share their high-net-worth investor sources in the syndicate during the process of book building and road shows. Book building is a practice by which bookrunner collects confidentially bids from institutional holders at various prices and disseminates noisy information about the issuer to these holders. As such, one can expect that a more central bookrunner in its underwriter network can potentially induce more institutional holders to buy shares from the issuer.

The objective of this paper is to analyze how the central position of a bookrunner in its underwriter network affects the position of an issuer in its public firm network in IPO process, and to study how the established position of the issuer in public firm network influences its post-IPO market performance. In the existing literature, effect has been seen of underwriter network on firm IPO performance (Bajo et al. (2016)). Nevertheless, the mechanism of underwriter network on firm performance is unclear. To further investigate the mechanism behind the effect, I study two research questions. The first question I analyze is how underwriter network affects firm network through IPO process. During book running process, bookrunner can induce institutional holders to pay attention to the firms it takes public by delivering noisy information to and collecting confidential bids from institutional holders. Intuitively, a bookrunner with higher network centrality in underwriter network has more sources on institutional holders, and can thus attract more investor attention, leading to higher network centrality of the issuer it takes public. Second, I examine how firm network affects issuer post-IPO market performance. Such effect could exist because institutional holders can affect firm's management decision and financial policy by setting board members. Moreover, institutional holders can deliver noisy information related to the firm to other firms through actions they take, such as selling a firm's shares or buying new shares.

To empirically address the two questions, I first construct firm and underwriter network, then use reduced form analysis. To construct the two networks, I use data from Thomson Reuters 13f institutional holdings and Thomson Reuters SDC New Issues Database. To capture firm performance, IPO characteristics, and firm characteristics, data from Center for Research in Security Prices (CRSP) and Compustat are used. Firm network is constructed by defining the linkages between two firms based on if they share a same institutional holder. Following the literature, I define underwriter network in the way that two underwriters are linked if they ever cooperated in a same syndicate in the previous five years before IPO. My dependent variables are three post-IPO stock market performance measures: average monthly trading volume, holding period return in the first post-IPO year and holding period return in the third post-IPO year. My controls are variables that capture firm characteristics, IPO characteristics, and market environments. I adopt two regression models of network centrality measures as independent variables. The centrality measures are degree, eigenvector, closeness, and betweenness, from social network analysis (SNA). The first model is a linear regression of firm network centralities on underwriter networks centralities. The second model is a linear regression of firm post-IPO performance on firm network centralities.

Empirical findings indicate that an issuer is significantly more central in its public firm network if this issuer is led by a bookrunner through IPO process that is more central in its underwriter network. For example, if a bookrunner has 40 more direct links with other underwriters, the firm to be issued will have about 25 more links with other public firms. I also find that the effect of firm network on issuer post-IPO market performance is significant and twofold. For example, if issuer betweenness centrality increases from minimum value to median value in the sample, the average monthly trading volume will significantly increase by around 7%. However, if issuer eigenvector centrality increases, the holding period return in the third year will significantly drop. To further investigate the source of the effect of underwriter network on firm performance, I run a specification by adding underwriter network centralities to the analyses of firm post-IPO performance. Empirical results show that firm network still has significant effects on firm's post-IPO stock market performance. Furthermore, the scales of effects that come from firm network is much larger than that that come from underwriter network. This finding indicates that the effect of underwriter network on firm performance that is documented in literature could be attributed to the effect of underwriter network on firm network.

This paper contributes to the existing literature of network effect in several aspects. First, I extend underwriter network literature by showing that firm network is more powerful in predicting stock market performance. In existing literature, underwriter network has been seen as an important predictor for issuer post-IPO performance (Bajo et al. (2016)). I show that the effect of underwriter network on firm performance is generated by the effect of underwriter network on firm network. Furthermore, it has been widely argued that favorable positions are regarded as network resources (Granovetter (1985), Snehota and Hakansson (1995), Gulati (1999), Burt (2000)). However, there are also abundant works on arguing that over-embeddedness in network can lead to inability to act (Uzzi (1997), Gargiulo and Benassi (2000)). My results also show the two-sided effect of being at favorable positions. An issuer with higher closeness centrality will achieve higher holding period return in the first post-IPO year, while an issuer with higher eigenvector centrality will have lower holding period return in the third post-IPO year.

Second, I construct inter-firm network in a novel way that based on common large institutional holders. By defining the linkages between two firms based on whether they share a same institutional holder, I show the key roles played by institutional holders in the relation between established firm network and firm stock market performance. Institutional holders can affect firm's management decision and financial policy by setting board members. Also, institutional holders can deliver information that are related to the firm to other firms in the network through actions they take, such as selling huge amount of shares of the firm. My work adds a new dimension to the existing business network studies. In the related literature, firm networks are defined in various ways. Previous research on inter-firm network has focused on firm linkages as mergers, acquisitions, R&D alliances, customers and suppliers, and so forth. Some researchers study firm networks in the concept of social networks where relationships with other firms are based on strong personal relationships with individuals such as friends, relatives, long-standing colleagues and so forth (Dubini and Aldrich (1991), Lipparini and Sobrero (1997), Ardichvili, Cardozo and Ray (2003),Lechner, Dowling and Welpe (2006), Cooper (2017)). There are also works on reputational networks that made up of partner firms that are market leaders (Lechner, Dowling and

Welpe (2006), Dietz (2012)), and co-operative technology networks that technology alliances involve joint research and development or innovation projects(Deeds and Hill (1996), Kelley and Rice (2002), Lechner, Dowling and Welpe (2006)). In recent literature, many researchers are interested in studying the network structure of customer-supplier relationships (Bengtsson and Kock (2000), Cohen and Frazzini (2008), Ahern and Harford (2014)).

Third, this is one of the first papers to study how one network can affect another network with totally different agents, to the best of the author's knowledge. Such study is important because network formation could be partly driven by other network as well as the strategic decisions made by agents in the network per se. I study this relationship by analyzing the impact of various network centrality measures characterizing bookrunners on network centrality measures characterizing issuers. I show that bookrunner centralities have significantly positive effects on issuer centralities. Many works in the literature study the relation between network structure or network centrality with network outcomes (Hanna and Walsh (2002)), Pappas and Wooldridge (2007), El-Khatib, Fogel and Jandik (2015)). For example, industries that are more central in the network of intersectoral trade earn higher stock returns than industries that are less central (Ahern (2013)). Firm that goes public with a lead underwriter that is more central in its investment bank network performs better (Bajo et al. (2016)). However, little work has been done on analyzing the relation between one network structure with another network structure. My study fills in the gap.

Last, this paper potentially provides a new way to capture investor attention through the construction of firm network. The linkage between two agents in the established firm network is constructed in terms of whether the two agents share a same holder. A firm with high degree centrality or eigenvector centrality is very likely to be held by many institutional holders or held by few holders that hold many other firms' shares, meaning that this firm gains a lot of investor attention. As such, the degree centrality and eigenvector centrality can be two good proxies for investor attention. Investor attention and investor inattention are two topics that have been discussed and followed for decades (Hong and Stein (1999), Huberman and Regev (2001), Hirshleifer and Teoh

(2003), Peng and Xiong (2006)). In past studies, investor attention is characterized using proxies, such as pre-IPO media coverage (Liu, Sherman and Zhang (2014), Bajo et al. (2016)), past news and extreme past returns. This paper potentially provides a new way to capture investor attention through the construction of firm network. Furthermore, under my framework, with annually updated data on institutional holdings, it is possible to get proxies for investor attention over time.

The rest of the paper is organized as follows. Section 2.1 describes the data I use and presents descriptive statistics on dependent variables and control variables. Section 2.2 introduces the four measures, degree, eigenvector, closeness and betweenness, that characterize network centralities and basic regression models. Section 2.3 discusses my empirical results and specifications. Section 2.4 discusses conclusions and future works.

2.1 Data

In this section, I introduce the data as well as the sample selection rules. The data comes from four main data sources: Thomson Reuters 13f institutional holdings, Thomson Reuters SDC New Issues Database, Compustat and Center of Research on Security Prices (CRSP).

The data on initial public offerings, are from the Thomson Reuters SDC New Issues Database over the period January 1987 to December 2009. Following the literature, I exclude firms in the financial sector with SIC code start from 6000 to 6999. Only issuances on new common equities that processed in U.S. market are kept, limited to three exchange markets: NASDAQ, AMEX and NYSE. Besides, following IPO literature, I exclude certain types of security on trust Units, units, non-voting shares and spin-offs. Eventually, I get a sample of IPO observations with size 3984. For each initial public offering, I observe the issuer (known as the firm that goes public), the bookrunner (also known as the lead manager or lead underwriter in the issuance of new equity) and the co-managers (that follow the IPO process along with the bookrunner in a syndicate). Following the literature, the underwriter network for each issuance observation is constructed by the chance of cooperation in the same syndicate within five years. For example, if manager A and manager

B worked in the same syndicate within five years prior to the IPO issue date, then A and B are considered connected.

The data on firm network are from the Thomson Reuters Institutional Holdings (13f) Database over the period January 1987 to December 2009. Observations from Institutional Holdings Database are matched with observations from SDC New Issues Database by Ticker symbol and issue date. Eventually, 3078 issuer observations are matched to New Issue Database. For each record, I observe the firm's name and the firm's institutional shareholders. For simplicity, I only keep the records on December 31th each year, since institutional holdings won't change much within a year. For a calendar year, the firm network is constructed by the chance that two firms share a same institutional holder. For example, if firm A has shareholders a and b, firm B has shareholders a and c, firm C has shareholder c, then A and B, B and C are considered connected, but A and C are not.

The control variables and firm performance variables are from Thomson Reuters SDC New Issues Database, the Center for Research in Security Prices (CRSP) and Compustat. Observations from both CRSP and Compustat are matched with observations from SDC New Issues database by Ticker symbol and issue date. Since the observations from different sources are matched according to the same calendar year, the possibility of change in Ticker symbol will not be problematic.

Table 1: Summary Statistics: Control Variables

Variable	Mean	Std. Dev.	Min.	Max.	Sample Size
LnOffersize	19.774	1.049	15.607	24.502	3984
HiTechDummy	0.565	0.496	0	1	3984
LnPrimaryShares	15.043	0.764	11.695	18.789	3984
FilingWidth20Dummy	0.078	0.268	0	1	3984
MiddleFiling	0.133	0.554	0.009	16.667	2512
AvgUnderpricing	-3.739	5.071	-31.171	11.435	3984
BMktshare	0.026	0.024	0.001	0.115	3984
MktReturn	0.013	0.037	-0.225	0.128	3731
LnAssets	2.906	1.68	-5.809	8.835	1652
OIBD/Asset	0.323	36.976	-402.3	1269.538	1364

Summary statistics of control variables and firm performance variables are shown in Table 1 and Table 2, respectively. *LnOffersize* is the log of IPO issuing principal amount (in dollars). *HiTechDummy* is an indicator variable that equals one if the firm is a high technology firm. *LnPrimaryShares* is the log of the number of primary shares offered in the exchange market. *FilingWidth20Dummy* is a dummy variable that equals one if the firm has a filing width greater or equal to 20%. Here, filing width is defined as the ratio of the gap between high filing price and low filing price to high filing price. *MiddleFiling* is defined as a unit divided by the middle point of filing price range. *AvgUnderpicing* is the average underpricing of all issuances in the prior month before the issue date of the underlying firm, where underpricing is simply calculated as the difference between the offer price and closing price at the 1st trading day. *BMktshare* is the market share of a bookrunner, computed as the ratio of the issuance amounts that led by the same bookrunner in the prior five years to the total issuance amounts processed in the prior five years before issue date. *MktReturn* is the value-weighted return on CRSP index in the quarter of the firm's issue date, which captures the market environment of an IPO. *LnAssets* is the log of total assets of a firm in the last quarter of the calendar year prior to the IPO issue date, which is a measure of firm size. *OIBD/Asset* is the firm's operating income before dividends to total assets in the last quarter of the calendar year before the IPO issue date. *LnOffersize*, *HiTechDummy*, *LnPrimaryShares*, *FilingWidth20Dummy*, *1/MiddleFiling*, *AvgUnderpicing* and *BMktshare* are computed by using SDC New Issues Database. *MktReturn* is computed from CRSP Database. *LnAssets* and *OIBD/AssetAdj* are calculated based on the data I get from Compustat.

Table 2: Summary Statistics: Firm Post IPO performance

Variable	**Mean**	**Std. Dev.**	**Min.**	**Max.**	**Sample Size**
LnTurnover	0.145	0.89	-4.943	3.633	3906
adj_HPR1	0.176	0.961	-0.971	12.828	3764
adj_HPR3	0.253	1.447	-1.24	37.735	3085

In the literature, there are various ways to capture firm post-IPO stock market performance. In this paper, I use *LnTurnover*, *HPR1*, and *HPR3* as the three market-based performance measures that obtained from CRSP dataset. *LnTurnover* is the log of average monthly share volumes traded in the first post-IPO year. *HPR1* and *HPR3* are the holding period return of the issuer's stock in the first post-IPO year and in the third post-IPO year, respectively.

2.2 Model

In this section, I introduce the centrality measures that characterize the position of an agent in one network. There are four measures from social network analysis (SNA) that are included in my models: degree, eigenvector, closeness and betweenness. To address the first research question, I analyze a linear regression model with bookrunner centrality measures as independent variables and firm centrality measures as dependent variables. To address the second research question, I adopt a linear regression model with firm centrality as independent variables and firm stock market performance as dependent variables.

2.2.1 Measures Characterizing Network Centrality

In this paper, a network consists of two components: vertices that represents agents and edges that represents the linkages between agents. Especially, an ego is used to represent the agent whose centrality importance is being analyzed, and an alter to represent the agents who are directly connected to the ego. The network can be either directed or undirected, depending on whether I distinguish the two connected vertices. In this paper, I only consider undirected network, meaning there is no distinction between the two vertices associated with each edge. Besides, I only consider unweighted network where all the edges are equally important.

Before measuring the importance of a vertex in a network, an adjacency matrix is needed. In an adjacency matrix, the ij-th element in the matrix equals to 1 if the i-th agent is connected with the j-th agent, and equals to 0 if they are not connected. Based on the corresponding adjacency matrix, I can calculate several measures that reflect the vertex importance in the network. These measures are known as network centrality, which are widely used in the field of social network analysis (SNA). In this paper, I use four different centrality measures: degree, eigenvector, closeness, and betweenness. Degree centrality and eigenvector centrality are the two centrality measures that are constructed based on direct links to the ego; while closeness centrality and betweenness centrality are the two measures obtained based on all the direct and indirect links to the ego. Because of the difference, degree centrality and eigenvector centrality are good measures for characterizing local importance of the ego; while closeness centrality and eigenvector centrality are used for capturing global importance of the ego. Details on these four measures are introduced in the following.

Degree is the simplest and most straightforward centrality measure that counts the number of edges directly connected to the ego. For a network with N vertices and adjacency matrix A, the degree of ego i, denoted as d_i, is represented mathematically as:

$$d_i' = \sum_{j \neq i}^{N} a_{ij}$$

Degree can be interpreted as the immediate risk for a vertex to catch information that flows through a network. For example, an ego with a larger degree centrality is exposed to more pathways that linked to its alters, hence the ego catches more information from its alters. To enable comparison on the degree centralities of two vertices from networks with different sizes, degree centrality is usually normalized by dividing the total number of potential direct links to the ego, $N-1$.

$$d_i = \frac{1}{N-1} \sum_{j \neq i}^{N} a_{ij} \tag{2.1}$$

Eigenvector is another measure that characterizes the local centrality of a vertex. Different from degree centrality that treats every direct connections to the ego equally, eigenvector highlights the differences in influence level of each connections. This measure assigns relative scores to all vertices in the network so that high-scoring vertices are more influential than low-scoring vertices. Then a high eigenvector for a vertex indicates that this vertex is connected to alters that are well-connected. The assigned score for a vertex i is given in the following way:

$$e_i = \frac{1}{\lambda} \sum_{j \in N(i)} e_j$$

, where $N(i)$ is the set of all the alters or neighbors to the ego i and λ is a constant. If the network G has an adjacency matrix $A = (a_{ij})$, then the assigned score can be written as:

$$e_i = \frac{1}{\lambda} \sum_{j \in G} a_{ij} e_j \tag{2.2}$$

If the centrality vector for the network is defined as $e = (e_1, e_2, ..., e_i, ..., e_N)'$, I can rewrite the above equation in a matrix form:

$$\lambda \cdot \mathbf{e} = \mathbf{A} \cdot \mathbf{e}$$

Clearly, $\mathbf{e}$ is an vector of matrix $\mathbf{A}$ with eigenvalue λ. Then by the Perron-Frobenius theorem that asserts a real square matrix with nonnegative entries has a unique largest real eigenvalue and that the corresponding eigenvector can be chosen to have nonnegative components. The normalized eigenvector corresponding to the unique largest eigenvalue of the adjacency matrix is defined as the eigenvector centrality.

Closeness is a centrality measure that is defined based on the concept of network paths, whereas degree centrality and eigenvector centrality are both defined based on direct connections. Closeness centrality tells us the extent to which a vertex is closer to all the other reachable vertices in a network either directly or indirectly. It is defined as the reciprocal of the average geodesic distance (known as the shortest path between a specified pair of vertices) between the ego and all of its reachable vertices:

$$c_i' = \frac{1}{\frac{\sum_{j \neq i} d(i,j)}{|R(i)|^2}} \in (0,1),$$

where $R(i)$ is the set of all reachable vertices of the ego i in the network and $|\cdot|$ gives us the cardinality of this set. Besides, $d(i,j)$ denotes the geodesic path between i and j.

To enable comparison between two networks with different sizes, the normalized expression of closeness centrality is given by:

$$c_i = \left(\frac{|R(i)|}{N-1}\right)^2 \frac{1}{\sum_{j \neq i} d(i,j)}. \tag{2.3}$$

According to the definition, it is clear that a higher closeness centrality of a vertex indicates a lower average distance of this vertex to all the other reachable vertices, hence the more central the vertex is. High closeness centrality can be interpreted as high chance of receiving information in a timely way.

Another measure that is defined based on the concept of network paths is betweenness centrality. This measure reflects how often an underlying vertex to occur on a randomly chosen shortest path between two randomly chosen vertices in a network. It is represented mathematically as:

$$b'_i = \sum_{j,k \neq i} \frac{n_{jk}(i)}{N_{jk}},$$

where N_{jk} is the total number of shortest paths from j to k and $n_{jk}(i)$ is the number of shortest paths from j to k that pass through the ego i.

Similarly, I need a normalized form of betweenness to enable comparison of two networks with different sizes, which is given in the following way:

$$b_i = \frac{b'_i}{\binom{N-1}{2}} = \frac{2b_i}{(N-1)(N-2)}. \tag{2.4}$$

A vertex with higher betweenness links together vertices who are otherwise unconnected or remotely connected, hence creating more chances to explore information or to influence the spread of information through the network.

2.2.2 Effect of One Network on Another Network

There are two networks in my model. One is the firm network, generally denoted as $G^f \equiv (V^f, E^f)$, where V^f is the set of all the firms that reported their institutional holders and E^f is the set of edges, which are 2-element subsets of $V^f \equiv (v_1^f, v_2^f, ...v_N^f)'$. The other is the underwriter network, denoted as $G^u \equiv (V^u, E^u)$, where $V^u \equiv (v_1^u, v_2^u, ...v_M^u)$ is the set of all underwriters that ever followed at least one IPO process during the underlying time period and E^u is the set of associated edges. Since I only focus on the position of the bookrunner (or the lead underwriter) in its underwriter network and in reality there are some underwriters never led an IPO, I need a subset of V^u: $V^B \subseteq V^u$. Similarly, I study the position of the issuers in the firm network, hence a subset of V^f: $V^I \subseteq V^f$ is needed. The crucial step that connects firm network and underwriter network is initial public offering (IPO). An IPO process is is defined as:

$$IPO_k = (t_k, v_k^I, v_k^B, V_k^f, V_k^u, H_k^I), \quad k \in \{1, 2, ..., K\}, \tag{2.5}$$

where t_k is the IPO issue year, v_k^I is the so called issuer or the firm that went public and v_k^B is the bookrunner or lead underwriter that led the IPO process for firm v_k^I. V_k^u is the set of all underwriters that worked for issuer v_k^I (known as the syndicate). V_k^f is the set of all firms that reported their institutional holders during the same issue year. Lastly, H represents the set of institutional holders that held a firm's stock shares. Specifically, H_k^I is the set of institutional holders that held large amount of common stock shares of firm v_k^I at the end of its issue year.

The formal definitions of the two networks and network measures are given as below:

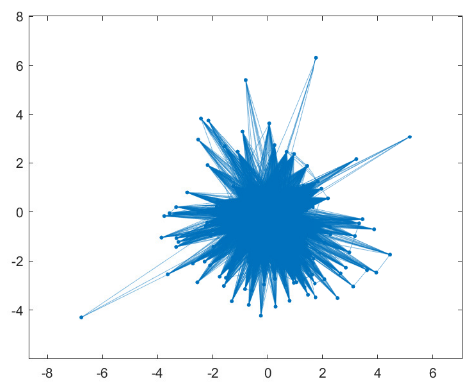
Definition 2.1 (Firm Network)**.** *In the k-th IPO year t_k, the firm network is defined as,*

$$G_k^f = (V_k^f, E_k^f), \tag{2.6}$$

where $V_k^f = (v_{1k}^f, ..., v_{ik}^f, ..., v_{n_k k}^f)$ is the set of all firms that reported their shareholdings in the *calendar year t_k. Let $A_k^f = (a_{ijk}^f)$, $i, j \in \{1, 2, ..., n_k\}$, be the adjacency matrix of G_k^f, where (1) $a_{ijk}^f = 1$ if $i \neq j$ and $H_{ik} \cap H_{jk} \neq \varnothing$, (2) $a_{ijk}^f = 0$ otherwise.*

Here, firm network is defined according to each calendar year. The total set of agents in firm network is the set of firms that reported their institutional holdings. The linkages of two firms are defined based on if they share a same institutional holder in that specific calendar year. For example, there are 3078 firms in the network in 2000 and the network can be plotted as in Figure 2.

Figure 2: Firm Network in Year 2000



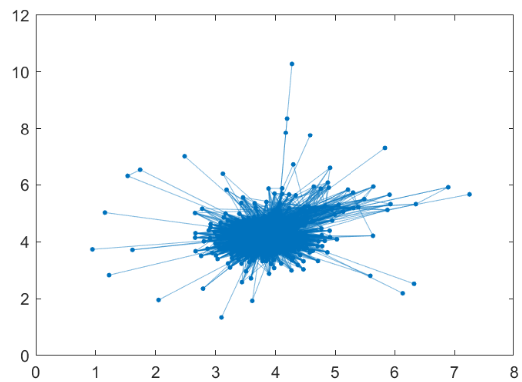
Definition 2.2 (Underwriter Network)**.** *In the k-th IPO year t_k, the underwriter network is defined as,*

$$G_k^u = (V_k^{*u}, E_k^{*u}), \tag{2.7}$$

where $V_k^{*u} \equiv \bigcup_{\tilde{k}: t_k - 5 \leqslant t_{\tilde{k}} \leqslant t_k} V_{\tilde{k}}^u$, interpreted as the set of underwriters that ever followed an IPO process either as bookrunners or co-managers within the five years prior to t_k. Let $A_k^u = (a_{ijk}^u)$, $i, j \in \{1, 2, ..., m_k\}$, be the adjacency matrix of G_k^u, where (1) $a_{ijk}^u = 1$ if $\exists$ some $\tilde{k}$ such that $t_k - 5 \leqslant t_{\tilde{k}} \leqslant t_k$, $i \neq j$, $v_{ik}^u \in V_{\tilde{k}}^u$ and $v_{jk}^u \in V_{\tilde{k}}^u$, (2) $a_{ijk}^u = 0$ otherwise.

Here, underwriter network is defined according to each IPO observation. For a specific date of IPO process, the linkage of two underwriters in the underwriter network is defined based on if the two underwriters had worked in the same syndicate in the previous five years before this IPO date. For example, the underwriter network that was constructed under the IPO of Neoforma.com INC on January 24th, 2000, can be plotted as shown in Figure 3.

Figure 3: Underwriter Network in Year 2000



Definition 2.3 (Network Measures)**.** *In the k-th IPO, the measures that characterizing the importance of issuer v_k^I in its network G_k^f is denoted as $M(v_k^I)$:*

$$M(v_k^I) = (d(v_k^I), e(v_k^I), c(v_k^I), b(v_k^I))'.$$

For simplicity, rewrite $M(v_k^I)$ as M_k^I:

$$M_k^I = (d_k^I, e_k^I, c_k^I, b_k^I)'.$$

In the k-th IPO, the measures that characterizing the importance of bookrunner v_k^B in its network G_k^u is denoted as $M(v_k^B)$:

$$M(v_k^B) = (d(v_k^B), e(v_k^B), c(v_k^B), b(v_k^B))'.$$

For simplicity, rewrite $M(v_k^B)$ as M_k^B:

$$M_k^B = (d_k^B, e_k^B, c_k^B, b_k^B)'.$$

In Definition 3, $d(\cdot)$, $e(\cdot)$, $c(\cdot)$, and $b(\cdot)$ are the degree centrality measure, the eigenvector centrality measure, the closeness centrality measure, and the betweenness centrality measure, respectively.

To study the relationship between firm network and underwriter network, I conduct a linear regression model (in matrix forms):

$$M^I = \delta_0 + \delta_1 M^B + u, \tag{2.8}$$

where

$$\delta_1 = \begin{bmatrix} \delta_d & 0 & 0 & 0 \\ 0 & \delta_e & 0 & 0 \\ 0 & 0 & \delta_c & 0 \\ 0 & 0 & 0 & \delta_b \end{bmatrix}.$$

2.2.3 Effect of Network on Firm Performance

I am also interested in studying how firm network affects firm post-IPO performance as well as the relationship between the two networks. To accomplish this goal, I conduct a linear regression model as follow:

$$y = \alpha_0 + M^{I'}\alpha_1 + C'\alpha_2 + \varepsilon_1, \tag{2.9}$$

where C is a vector of control variables.

$$\alpha_1 = (\alpha_{1d}, \alpha_{1e}, \alpha_{1c}, \alpha_{1b})'.$$

A specification regression is conducted to show the robustness of regression results from equation (5).:

$$y = \beta_0 + M^{I'}\beta_1 + M^{B'}\beta_2 + C'\beta_3 + \varepsilon_2, \tag{2.10}$$

where

$$\beta_1 = (\beta_{1d}, \beta_{1e}, \beta_{1c}, \beta_{1b})', \quad \beta_2 = (\beta_{2d}, \beta_{2e}, \beta_{2c}, \beta_{2b})'.$$

2.3 Empirical Results

In this section, I present my main empirical results based on the methodology that was introduced in Section 2.2. Table 3 and Table 4 report the summary statistics of both issuer centralities and bookrunner centralities in our regression analyses in subsequent sections. Table 5 presents the regression results of issuer network centralities on bookrunner network centralities. It shows that all the four centrality measures of bookrunners have significantly positive effects on the four measures of issuers, meaning that a bookrunner with higher centralities in its underwriter network will help the associated issuer obtain higher centralities as well in the firm network. If I control for firm size, firm operating ability, and some other characteristics, the underwriter centrality measures still have significant effect on firm centrality measures, except for eigenvector. Table 6 to Table 8 show the main regression results of issuer Post IPO performances on issuer network centralities. The results show that issuer network centralities do have some significantly positive impacts on the issuer performances after they go public. Especially, the most predictable measure is betweenness centrality, which characterizes an issuer's global importance and the ability an issuer can catch or control the information that pass through firm network. Table 10 to Table 11 present the specification studies on the regression of issuer post IPO performances on both issuer centralities and bookrunner centralities. From these specifications, conclusions are drawn that the impact of bookrunner centralities on issuer performance is much less than the impact of issuer centralities on issuer performance. Details on interpretation of these regression results are given in the subsequent subsections.

2.3.1 Network Centralities Measures

The descriptive statistics presented in Table 3 summarize the network centrality measures of issuers in their own firm networks. I use *Fdegree*, *Feigenvector*, *Fcloseness* and *Fbetweenness* to stand for each issuer's degree centrality, eigenvector centrality, closeness centrality, and betweenness centrality, respectively. Following Definition 1, the firm networks are obtained by calendar year, leading to 23 different networks since there are 23 years in the period of January 1987 to December 2009. As a result, I observe different network sizes for different firm networks. To enable comparison on centralities from different networks, all the measures are normalized.

Table 3: Summary Statistics: Firm Network Measures

Variable	**Mean**	**Median**	**Std. Dev.**	**Min.**	**Max.**	**Sample Size**
Fdegree	0.777386	0.872529	0.232342	0.002933	0.985108	3078
Fcloseness	0.000361	0.000333	0.000111	0.000161	0.000662	3078
Fbetweenness	0.000038	0.000030	0.000037	0	0.000551	3078
Feigenvector	0.000399	0.000372	0.000163	0.000001	0.000790	3078

The descriptive statistics shown in Table 4 summarize the network centrality measures of bookrunners in their own underwriter networks.I use *Udegree*, *Ueigenvector*, *Ucloseness* and *Ubetweenness* to stand for each observed bookrunner's degree centrality, eigenvector centrality, closeness centrality, and betweenness centrality, respectively. I also need to normalize all the four measures because it is clear that the underwriter network sizes also vary across observations, by Definition 2.

Table 4: Summary Statistics: Underwriter Network Measures

Variable	**Mean**	**Median**	**Std. Dev.**	**Min.**	**Max.**	**Sample Size**
Udegree	0.277028	0.241191	0.262007	0	0.774389	3984
Ucloseness	0.000126	0.000129	0.000117	0	0.000415	3984
Ubetweenness	0.003907	0.000719	0.005712	0	0.047903	3984
Ueigenvector	0.001703	0.001391	0.000796	0	0.003409	3984

The data on IPO records are from SDC New Issue Dataset from January 1982 to Decemeber

2009, with sample size 4788. For each IPO record, I observe one issuer and its associated bookrunner and co-managers. Based on these records, I can construct underwriter networks in the period of January 1987 to Decemeber 2009, with sample size 3984. Unfortunately, not all IPO issuers are recorded in the Thomson Reuters 13f dataset that reveals their institutional holdings. Eventually, 3078 IPO issuers are matched with the 13f dataset, hence 3078 issuer centrality measures are observed.

Interesting patterns are revealed if I plot issuer centralities and bookrunner centralities in cumulative way, as shown in Figure 4. The degree centralities of issuers are larger than the degree centralities of bookrunners at every cumulative density level. For example, at the 60% quantile, issuer centrality is around 0.9 while bookrunner centrality is only about 0.4. Especially, there are almost 35% of bookrunners are isolated in their own networks, with 0 degree centralities, meaning they did not join any syndicates with other underwriters in the past five years before it led an IPO. Besides, there are some issuers that have degree close to 1, meaning they are connected to almost every other firms in the network. If I examine the adjacency matrices that associated with the networks, I observe dense matrices for firm network and relatively sparser matrices for underwriter network. The difference in the density level come from the difference in defining the two networks. As for closeness centrality, I observe a similar pattern to degree centrality that the closeness centralities of issuers are larger than the closeness centralities of bookrunners at every cumulative density level. The closeness centrality measures the average distance of an ego to all the other agents it can reach, hence it can tell us to what extent an ego is close to other agents. On average, the issuers are closer to other firms in their firm networks compared with the bookrunners in their underwriter networks.

It is less meaningful to compare the scattered cumulative density diagram of eigenvector between issuers and bookrunners, due to normalization. To get a unique corresponding eigenvector for an adjacency matrix, the eigenvector is normalized so that the sum of all components is 1. If two networks exhibit quite a difference in size, the eigenvectors will differ a lot. In this paper, the

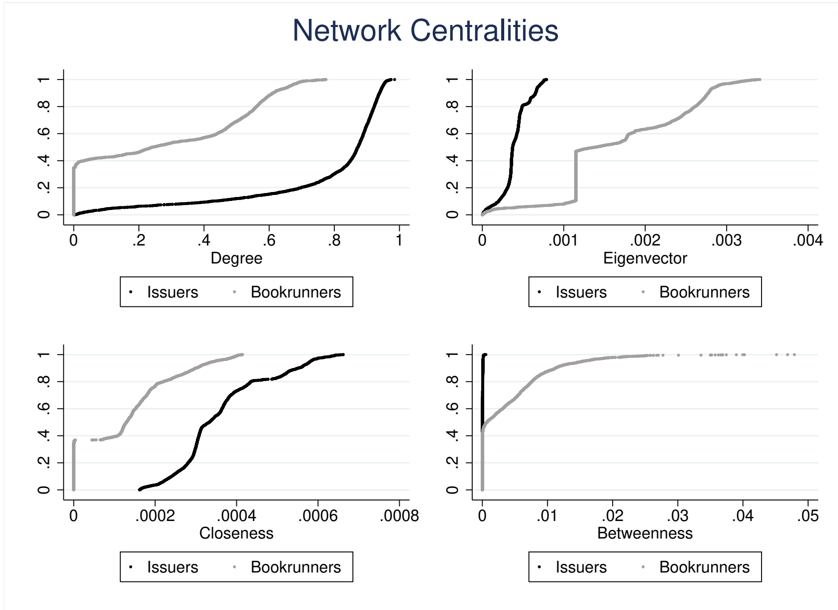


Figure 4: Network Centralities

firm network has a size of 3000 on average while the underwriter network has a size of 400 on average. So when it comes to normalized eigenvector, clearly I should observe much less scales in issuer eigenvector centralities. Similar scale issue happens to betweenness centrality, too. Betweenness is normalized by dividing $(N-1)(N-2)$, leading to huge scale differences between the issuer betweenness centralities and the bookrunner betweenness centralities. Even though certain scale issue exists, our linear regression model is still valid as long as careful examinations on the scale of each predicted coefficient are conducted.

2.3.2 Firm and Underwriter Network Centralities

Table 5 presents the results from the four regression models in equation (2.10) with year dummies. These regression results indicate, as expected, that the centralities of bookrunners have significant positive effects on the centralities of issuers that go public. The intuition behind this significant positive effects lies in book building, a process in which bookrunner attempts to determine offer price for IPO based on the demand information they get from institutional holders. During the book building process, it is true that a bookrunner will contact some institutional holders that potentially will buy the new issuer's stocks. If a bookrunner is in a central position of its underwriter network, then it will have more sources on the institutional holders. Then the pool of institutional holders the central bookrunner can contact to is obviously larger than that a remote bookrunner can contact to. The possibility that the central bookrunner can persuade or lead institutional holders to buy the issuer's stock is higher than a remote bookrunner can do. As such, a bookrunner that is central in its underwriter network will help the issuer that goes pubic build more direct links to other firms or build closer relationship to other firms.

The coefficient per se carries useless information on the scale of effect unless the difference between the firm network size and underwriter network size are taken considered. In the sample, most firm networks have size around 3000, while most underwriter networks have size around 400. To interpret 0.0842, the coefficient of bookrunner degree, I should scale up the network size. For example, if a bookrunner has 0.1 higher degree centrality, meaning the bookrunner connects to 40 more other underwriters, then the corresponding issuer will have 0.00842 higher degree centrality, meaning the issuer will connect to about 25 more public firms.

Table 5: Regression Results: Effect of Underwriter Centrality on Firm Centrality

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Udegree	0.0847***				0.0501**			
	(0.0157)				(0.0246)			
Ueigenvector		0.0140***				0.00453		
		(0.00307)				(0.00462)		
Ucloseness			0.0399***				0.0253*	
			(0.0107)				(0.0147)	
Ubetweenness				0.000559***				0.000489***
				(0.000116)				(0.000176)
LnAssets					0.0148***	0.00000784***	0.00000457***	0.00000340***
					(0.00440)	(0.00000235)	(0.00000113)	(0.000000619)
OIBD/Asset					0.0000767	3.74e-08	1.93e-08	1.15e-08
					(0.000153)	(8.15e-08)	(3.96e-08)	(2.15e-08)
HiTechDummy					0.0254	0.0000139*	0.00000783*	-0.000000199
					(0.0157)	(0.00000837)	(0.00000406)	(0.00000221)
AvgUnderpricing					-0.000837	5.39e-09	2.04e-08	-9.54e-08
					(0.00213)	(0.00000113)	(0.000000550)	(0.000000299)
MktReturn					-0.108	0.0000189	0.0000279	-0.0000160
					(0.183)	(0.0000975)	(0.0000473)	(0.0000257)
YearDummy	YES	YES	YES	YES	YES	YES	YES	YES
IndustryDummy	YES	YES	YES	YES	YES	YES	YES	YES
Constant	0.715***	0.000565***	0.000501***	0.0000618***	0.660***	0.000550***	0.000498***	0.0000485***
	(0.0313)	(0.0000181)	(0.00000871)	(0.00000493)	(0.0586)	(0.0000317)	(0.0000151)	(0.00000821)
Sample Size	3078	3078	3078	3078	992	992	992	992
R^2	0.069	0.390	0.683	0.107	0.095	0.437	0.717	0.190
adj. R^2	0.061	0.385	0.680	0.098	0.065	0.419	0.707	0.163

Standard errors in parentheses

Dependent variable of Column 1 and 5 is firm degree centrality;Dependent variable of Column 2 and 6 is firm eigenvector centrality;Dependent variable of Column 3 and 7 is firm closeness centrality;Dependent variable of Column 4 and 8 is firm betweenness centrality.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

2.3.3 Firm Performance and Network Centralities

In this subsection, I study the relation between issuer centrality and five issuer post-IPO performance measures. The three dependent variables are stock turnover rate, holding period return in the first post-IPO year and holding period return in the third post-IPO year.

The independent variables in the regressions are issuer degree centrality, issuer eigenvector centrality and issuer betweenness centrality. The closeness centrality is omitted because it has a very strong co-linearity with eigenvector centrality. Two specifications are given by dividing these four measures into two groups. One group contains degree and eigenvector, which are the two measures that characterize the local centrality level of an issuer. The other group includes closeness and betweenness, which are the two measures that are defined based on paths, hence telling the global centrality level of an issuer. The regression results are presented in Table 6 to Table 8. Sepecification results are shown in Table 9 to Table 11.

To compare the effect on issuer performance driven by issuer centralities with that driven by bookrunner centralities, I conduct specification regressions with both issuer centralities and bookrunner centralities as independent variables. Similar specification on local centralities and global centralities are conducted too. The regression results can be found in Table 10 to Table 11.

Clearly, besides the issuer network effect on issuer performance, issuer's IPO characteristics may also have non-negligible effects on the issuer post-IPO performance. I control for offer size by including natural logarithm of IPO total principal amount (*LnOffersize*). I also control for amount of primary shares listed in the primary exchange market(*LnPrimaryShares*). Both offer size and primary shares are likely to affect the post-IPO performances. Besides, I include high-tech issuer dummy as a control variable. High-tech issuers are expected to have higher valuations and yield better performance.

The uncertainty about the value of IPO shares to be issued is given by the dummy variable on whether the filing width is greater than 20% (*FilingWidth20Dummy*). Filing width is defined as the percentage difference between high filing price and low filing price that reported by bookrunner.

I also control for reciprocal of the middle point of filing range, which capture the effect of the choice of price level in IPO (*MiddleFiling*). I expect bookrunner that has higher market shares will be better at leading an IPO process, hence leading to issuer's better performance. So the market share of bookrunner is also included as a control variable (*BMktshare*).

Besides, market environment at the time issuer goes public should be controlled. The first control variable I use is the average underpricing in the prior month before IPO issue date(*AvgUnderpricing*). This variable capture recent IPO environment by telling us whether the market is a heating market with higher underpricing or not. The second control variable I sue is the CRSP value weighted market return in the prior quarter before IPO (*Mktreturn*).

Furthermore, I add year dummies in every regression to control for year effects. As for industry dummies, I conduct specification on whether industry dummies are added.

I study the relation between issuer centrality and issuer stock turnover rate by running regressions with issuer's average monthly trading volume in the first calendar post-IPO year as the dependent variable.

The main regression results are shown in column 1 and column 2 of Table 6. Among the centrality measures, only betweenness centrality is statistically significant and the impact of betweenness on turnover rate is positive. The estimated coefficients exhibit little difference if industry dummies are not included. The two groups specification results are shown in column 3 to column 6. In the specification that I only include global centrality measures, betweenness centrality still significantly and positively affect turnover rate, as expected.

Higher issuer betweenness centrality means that the issuer will occur more frequently on the shortest paths of other pairs of firms. By lying on the shortest paths between others, an issuer along with its institutional holders will catch information more efficiently and even be able to control information diffusion. Here, the institutional holders play roles as signals of such information by taking actions of selling issuer's stocks or buying issuer's stocks. Therefore, there will be more frequent transactions on the stocks of the issuer.

For example, consider two firms A and B, between which there are only four shortest paths with length 4. Firm X and Y are the two firms that occurs on the shortest paths, with betweenness $\frac{3}{4}$ and $\frac{2}{4}$, respectively. This means that firm X occurs three times on the four shortest paths, while firm Y occurs two times on the four shortest paths. Suppose I hold other measures the same for firm X and Y. If anything bad or good happened to firm A, the institutional holders of firm A will directly take actions, either selling the shares or buying more shares. Then the firms whose shares are held by those holders will also learn the fact happened to firm A. Step by step, the information will move along the shortest paths till firm B. Clearly, firm X with higher betweenness will be more likely to catch the information, even to manipulate the diffusion and its stocks will be more likely to be traded than firm Y.

Furthermore, offer size and bookrunner market share have significantly positive effects on turnover rate. Also, high-tech companies have significantly higher monthly average trading volume than non high-tech companies.

Table 11 presents the specification results of turnover rate on issuer and bookrunner centralities. It shows that the issuer betweenness is still positively significant. The bookrunner betweenness is also positively significant, but its effect on turnover rate is less than the issuer's. The effect magnitude is not straightforward if I only look at the coefficient because the scale of centrality measures is not straightforwardly comparable. Based on the sample I have, supposing both issuer betweenness and bookrunner betweenness change from minimum value to median value, the turnover rate will increase by 7% driven by issuer betweenness, but only 0.7% driven by bookrunner betweenness. These patterns tell us that, there are two ways that bookrunner affects turnover rate: directly or through affecting issuer network. Bookrunner centrality has very little direct effect on turnover rate. Instead, issuer centrality plays an important role in issuer turnover rate.

I study the relation between issuer centrality and issuer stock holding period return by running regressions with issuer's annualized holding period return in the first post-IPO calendar year as the dependent variable. Further, to show how this effect changes over time, I also run a regression with issuer's annualized holding period return in the third post-IPO calendar year as the dependent variable.

The first regression results that I am interested in lie in column 5 and column 6 in Table 7. The closeness centrality is significantly positive affecting the annualized industry-adjusted holding period return, when I regress adj_HPR1 only on global centrality measures. I did not get significant centrality measures in column 1 and 2 because I omit closeness which has co-linearity with eigenvector. The estimated coefficients exhibit little difference if industry dummies are not included.

The second regression results that I am interested in are shown in column 1 and column 2, Table 8. Here, dependent variable is the annualized industry-adjusted holding period return in the third post-IPO year. The closeness centrality is no longer significant and the effect is also diminished to negative values. But instead, the eigenvector is statistically significant and the impact of eigenvector on adj_HPR3 is negative. In the specification where I regress adj_HPR3 only on local centrality measures, same results still hold. Also, the estimated coefficients exhibit little difference if industry dummies are not included.

Higher issuer closeness means that the issuer is closer to all its reachable agents. One possible explanation why issuer closeness has positive effect on holding period return in the first post-IPO year is that, the closer the issuer to its reachable agents, the higher the possibility it is exposed to risks, hence yielding higher return by the modern portfolio theory.

When it comes to the holding period return in the third post-IPO year, closeness is no longer significant. Instead, eigenvector is negatively significant. Higher issuer eigenvector means that the issuer is connected to alters that are well connected in the network. One possible intuition is that, as time goes by, institutional holders' attention may gradually switch to those alters since they did well in the past two years. As such, it is likely that these institutional holders will sell many of the issuer's shares, leading to lower stock return.

Table 7 also shows that high-tech issuers have significant higher holding period return both in the first post-IPO year and in the third post-IPO year than the non-high-tech issuers. Bookrunner market share has a significantly positive effect on holding period return in the first post-IPO year, but not significantly on holding period return in the third post-IPO year. This pattern makes sense, since a bookrunner with higher market shares is likely to be better at leading IPO, leading to higher return of the issuer, but eventually the effect will drop once the bookrunner stops following the issuer. The filing width dummy exhibits a significantly positive effect on holding period return. One possible reason is that a higher uncertainty about the value of IPO shares to be issued implies a higher volatility in its stocks, hence possibly leading to higher return.

Table 10 and Table 11 show the specification results of holding period return on both issuer and bookrunner centralities. Table 10 shows that issuer closeness is still positively significant, while underwriter closeness is not significant at all. Instead, underwriter betweenness is positively significant but with a low magnitude. Table 11 shows that issuer eigenvector is still negatively significant, but none of the centrality measures of bookrunner is significant. These patterns tell us that, bookrunner centrality has very little direct effect on holding period return in the first post-IPO year, and no significant direct effect on the holding period return in the third post-IPO year. Instead, issuer centrality play an important role in issuer holding period return.

2.4 Conclusion

Network structure and network content have been seen as important factors for various aspects of network outcomes in previous studies on network effect. In this paper, I show that network structure can affect another network structure with totally different agents as well as certain network outcomes. I can draw two main conclusions by adopting two regression models of network centrality measures. First, an issuer is significantly more central in its public firm network if this issuer is led by a bookrunner through IPO process that is more central in its underwriter network. Second, the effect of firm network on issuer post-IPO performance is significant and twofold. Moreover, in the specification regression, I show that the effect of underwriter network on firm performance that was studied in literature is actually generated by the effect of underwriter network on firm network. There could be two feasible and enlightening tasks to be considered as part of future works. One can think of using weighted adjacency matrix instead of unweighted adjacency matrix to characterize firm network under the concept of institutional holdings. By weighting the adjacency matrix with different importance in linkages, one can tell whether a pair of firms are strongly connected because they share a same institutional holder that holds huge amount of shares in both firms. Also, one can think of using the degree centrality of a firm, either from a weighted adjacency matrix or an unweighted adjacency matrix, as a proxy for the ability in attracting investor attention of this firm. Since there is annually collected data on institutional holdings, it is possible to get proxies over time, hence to study how does investor attention could affect firm performance.

Table 6: Regression Results: Effect of Firm Centralities on Turnover Rate

	(1)	(2)	(3)	(4)	(5)	(6)
Fdegree	0.198	0.207	0.238	0.248		
	(0.204)	(0.205)	(0.204)	(0.205)		
Feigenvector	-505.2	-543.4	-323.2	-363.0		
	(349.8)	(350.8)	(346.9)	(347.8)		
Fcloseness					0.329	-90.59
					(327.0)	(327.3)
Fbetweenness	1974.5***	1951.8***			1672.5***	1684.9***
	(572.5)	(573.7)			(574.8)	(576.3)
LnOffersize	0.179***	0.182***	0.201***	0.203***	0.170***	0.174***
	(0.0503)	(0.0505)	(0.0501)	(0.0502)	(0.0503)	(0.0504)
HiTechDummy	0.340***	0.333***	0.338***	0.333***	0.340***	0.333***
	(0.0404)	(0.0374)	(0.0405)	(0.0375)	(0.0404)	(0.0374)
LnPrimaryShares	-0.260***	-0.269***	-0.259***	-0.268***	-0.251***	-0.261***
	(0.0691)	(0.0692)	(0.0693)	(0.0694)	(0.0690)	(0.0692)
FilingWidth20Dummy	0.0449	0.0438	0.0551	0.0538	0.0429	0.0418
	(0.0563)	(0.0565)	(0.0564)	(0.0566)	(0.0563)	(0.0565)
MiddleFiling	0.0721	0.0737	0.0746	0.0764*	0.0758*	0.0775*
	(0.0456)	(0.0458)	(0.0457)	(0.0459)	(0.0456)	(0.0457)
AvgUnderpricing	-0.00610	-0.00910	-0.00630	-0.00930	-0.00655	-0.00949
	(0.0105)	(0.0105)	(0.0105)	(0.0105)	(0.0105)	(0.0105)
BMktshare	4.142***	4.366***	4.090***	4.311***	4.109***	4.347***
	(1.052)	(1.051)	(1.055)	(1.054)	(1.052)	(1.051)
MktReturn	-0.1000	-0.0438	0.00897	0.0637	-0.107	-0.0495
	(0.556)	(0.558)	(0.557)	(0.558)	(0.556)	(0.558)
YearDummy	YES	YES	YES	YES	YES	YES
IndustryDummy	YES	NO	YES	NO	YES	NO
Constant	-1.005*	-0.648	-1.453***	-1.064**	-1.101**	-0.717
	(0.552)	(0.520)	(0.538)	(0.507)	(0.561)	(0.530)
Sample Size	1857	1857	1857	1857	1857	1857
R^2	0.208	0.200	0.203	0.195	0.207	0.199
adj. R^2	0.192	0.186	0.187	0.182	0.192	0.185

1. Dependent variable is LnTurnover, the monthly average trading volume in the first post-IPO year.

2. Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 7: Regression Results: Effect of Firm Centralities on HPR1

	(1)	(2)	(3)	(4)	(5)	(6)
Fdegree	-0.0487	-0.0326	-0.0334	-0.0172		
	(0.238)	(0.239)	(0.237)	(0.238)		
Feigenvector	240.4	197.7	309.8	266.7		
	(406.0)	(407.5)	(401.4)	(402.9)		
Fcloseness					688.1*	649.9*
					(377.1)	(377.8)
Fbetweenness	750.2	744.2			457.7	452.2
	(659.4)	(661.0)			(661.6)	(663.6)
LnOffersize	0.0335	0.0349	0.0423	0.0434	0.0289	0.0302
	(0.0580)	(0.0582)	(0.0575)	(0.0577)	(0.0579)	(0.0581)
HiTechDummy	0.0927**	0.139***	0.0919**	0.139***	0.0924**	0.138***
	(0.0465)	(0.0431)	(0.0465)	(0.0431)	(0.0464)	(0.0430)
LnPrimaryShares	-0.0344	-0.0406	-0.0346	-0.0409	-0.0312	-0.0372
	(0.0794)	(0.0796)	(0.0794)	(0.0796)	(0.0792)	(0.0795)
FilingWidth20Dummy	-0.0680	-0.0577	-0.0639	-0.0536	-0.0672	-0.0570
	(0.0648)	(0.0651)	(0.0648)	(0.0650)	(0.0648)	(0.0650)
MiddleFiling	0.0317	0.0320	0.0329	0.0332	0.0318	0.0323
	(0.0523)	(0.0525)	(0.0523)	(0.0525)	(0.0522)	(0.0524)
AvgUnderpricing	-0.00248	-0.00119	-0.00255	-0.00126	-0.00303	-0.00178
	(0.0122)	(0.0122)	(0.0122)	(0.0122)	(0.0122)	(0.0122)
BMktshare	2.269*	2.519**	2.249*	2.497**	2.193*	2.453**
	(1.206)	(1.206)	(1.206)	(1.206)	(1.204)	(1.204)
MktReturn	0.224	0.325	0.263	0.364	0.230	0.332
	(0.640)	(0.642)	(0.639)	(0.641)	(0.639)	(0.641)
YearDummy	YES	YES	YES	YES	YES	YES
IndustryDummy	YES	NO	YES	NO	YES	NO
Constant	0.155	0.180	-0.0178	0.0194	-0.0277	0.00216
	(0.635)	(0.600)	(0.617)	(0.583)	(0.645)	(0.610)
Sample Size	1834	1834	1834	1834	1834	1834
R^2	0.082	0.072	0.081	0.071	0.083	0.073
adj. R^2	0.064	0.056	0.063	0.056	0.065	0.057

1. Dependent variable is adj_HPR1, the industry-adjusted holding period return in the first post-IPO year.

2. Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 8: Regression Results: Effect of Firm Centralities on HPR3

	(1)	(2)	(3)	(4)	(5)	(6)
Fdegree	0.687	0.696	0.685	0.693		
	(0.454)	(0.454)	(0.454)	(0.454)		
Feigenvector	-1460.6*	-1483.3*	-1483.5*	-1519.1**		
	(778.4)	(778.4)	(766.6)	(766.6)		
Fcloseness					-724.9	-776.7
					(703.3)	(702.7)
Fbetweenness	-194.9	-304.8			-204.9	-285.0
	(1138.2)	(1136.2)			(1156.4)	(1155.3)
LnOffersize	-0.0229	-0.0105	-0.0251	-0.0139	-0.0316	-0.0191
	(0.106)	(0.106)	(0.105)	(0.105)	(0.106)	(0.106)
HiTechDummy	0.263***	0.237***	0.264***	0.237***	0.266***	0.241***
	(0.0836)	(0.0765)	(0.0835)	(0.0764)	(0.0835)	(0.0763)
LnPrimaryShares	0.0262	0.0101	0.0261	0.00998	0.0382	0.0218
	(0.144)	(0.144)	(0.144)	(0.144)	(0.144)	(0.144)
FilingWidth20Dummy	0.344***	0.353***	0.343***	0.351***	0.339***	0.348***
	(0.116)	(0.116)	(0.116)	(0.116)	(0.116)	(0.116)
MiddleFiling	-0.0922	-0.0911	-0.0926	-0.0918	-0.0891	-0.0879
	(0.102)	(0.102)	(0.102)	(0.102)	(0.102)	(0.102)
AvgUnderpricing	0.00760	0.00436	0.00759	0.00434	0.00700	0.00387
	(0.0230)	(0.0230)	(0.0230)	(0.0230)	(0.0231)	(0.0230)
BMktshare	0.740	1.263	0.741	1.266	0.873	1.404
	(2.125)	(2.118)	(2.124)	(2.117)	(2.125)	(2.117)
MktReturn	-0.775	-0.656	-0.786	-0.673	-0.801	-0.682
	(1.141)	(1.141)	(1.139)	(1.139)	(1.142)	(1.142)
YearDummy	YES	YES	YES	YES	YES	YES
IndustryDummy	YES	NO	YES	NO	YES	NO
Constant	0.0363	0.158	0.0840	0.228	0.0245	0.177
	(1.120)	(1.054)	(1.084)	(1.021)	(1.143)	(1.079)
Sample Size	1518	1518	1518	1518	1518	1518
R^2	0.083	0.079	0.083	0.079	0.082	0.077
adj. R^2	0.062	0.060	0.062	0.060	0.061	0.059

1. Dependent variable is adj_HPR3, the industry-adjusted holding period return in the third post-IPO year.
2. Standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 9: Specification Results: Effect of Firm Centralities on Turnover Rate

	(1)	(2)	(3)	(4)	(5)	(6)
Fdegree	0.190	0.199	0.227	0.235		
	(0.205)	(0.205)	(0.205)	(0.206)		
Feigenvector	-491.4	-529.0	-300.6	-339.1		
	(350.8)	(351.8)	(348.6)	(349.5)		
Fcloseness					3.168	-89.43
					(327.0)	(327.3)
Fbetweenness	1950.4***	1927.6***			1639.4***	1651.5***
	(572.2)	(573.2)			(574.1)	(575.6)
Udegree	-0.274*	-0.284*	-0.0982	-0.104		
	(0.158)	(0.158)	(0.146)	(0.146)		
Ueigenvector	42.30	44.83	32.41	35.27		
	(50.22)	(50.33)	(50.40)	(50.51)		
Ucloseness					-524.9**	-507.2*
					(264.9)	(265.8)
Ubetweenness	10.60***	10.91***			9.665***	9.778***
	(4.004)	(4.013)			(3.689)	(3.696)
LnOffersize	0.195***	0.198***	0.205***	0.208***	0.178***	0.182***
	(0.0511)	(0.0513)	(0.0508)	(0.0510)	(0.0505)	(0.0506)
HiTechDummy	0.343***	0.337***	0.339***	0.334***	0.341***	0.336***
	(0.0405)	(0.0375)	(0.0407)	(0.0376)	(0.0404)	(0.0374)
LnPrimaryShares	-0.276***	-0.285***	-0.265***	-0.274***	-0.258***	-0.267***
	(0.0699)	(0.0701)	(0.0701)	(0.0703)	(0.0692)	(0.0694)
FilingWidth20Dummy	0.0494	0.0486	0.0568	0.0557	0.0436	0.0426
	(0.0563)	(0.0565)	(0.0565)	(0.0566)	(0.0562)	(0.0564)
MiddleFiling	0.0758*	0.0775*	0.0752	0.0770*	0.0785*	0.0800*
	(0.0456)	(0.0458)	(0.0458)	(0.0460)	(0.0455)	(0.0457)
AvgUnderpricing	-0.00673	-0.00964	-0.00642	-0.00942	-0.00741	-0.0103
	(0.0105)	(0.0105)	(0.0105)	(0.0105)	(0.0105)	(0.0105)
BMktshare	3.627***	3.817***	4.063***	4.268***	3.649***	3.846***
	(1.159)	(1.160)	(1.150)	(1.151)	(1.121)	(1.122)
MktReturn	-0.108	-0.0534	0.0127	0.0681	-0.106	-0.0504
	(0.556)	(0.557)	(0.557)	(0.559)	(0.555)	(0.557)
YearDummy	YES	YES	YES	YES	YES	YES
IndustryDummy	YES	NO	YES	NO	YES	NO
Constant	-1.080*	-0.737	-1.478***	-1.094**	-1.136**	-0.758
	(0.552)	(0.522)	(0.539)	(0.509)	(0.560)	(0.530)
Sample Size	1857	1857	1857	1857	1857	1857
R^2	0.211	0.203	0.203	0.195	0.210	0.202
adj. R^2	0.194	0.189	0.187	0.181	0.194	0.188

1. Dependent variable is LnTurnover, the monthly average trading volume in the first post-IPO year.

2. Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 10: Specification Results: Effect of Firm Centralities on HPR1

	(1)	(2)	(3)	(4)	(5)	(6)
Fdegree	-0.0673	-0.0480	-0.0555	-0.0364		
	(0.238)	(0.239)	(0.238)	(0.239)		
Feigenvector	255.4	208.3	334.4	286.6		
	(407.3)	(408.8)	(403.0)	(404.5)		
Fcloseness					665.9*	628.7*
					(377.4)	(378.2)
Fbetweenness	750.0	740.0			428.8	424.3
	(659.2)	(660.9)			(661.4)	(663.5)
Udegree	-0.195	-0.177	-0.0788	-0.0645		
	(0.182)	(0.182)	(0.167)	(0.168)		
Ueigenvector	89.13	78.91	83.90	74.00		
	(57.72)	(57.89)	(57.69)	(57.86)		
Ucloseness					-155.5	-144.9
					(305.1)	(306.4)
Ubetweenness	7.206	7.020			8.274*	7.891*
	(4.587)	(4.601)			(4.230)	(4.239)
LnOffersize	0.0367	0.0379	0.0379	0.0391	0.0301	0.0311
	(0.0589)	(0.0591)	(0.0583)	(0.0585)	(0.0582)	(0.0583)
HiTechDummy	0.0895*	0.138***	0.0871*	0.135***	0.0908*	0.138***
	(0.0466)	(0.0432)	(0.0466)	(0.0432)	(0.0464)	(0.0430)
LnPrimaryShares	-0.0355	-0.0412	-0.0287	-0.0349	-0.0292	-0.0349
	(0.0804)	(0.0806)	(0.0803)	(0.0806)	(0.0796)	(0.0798)
FilingWidth20Dummy	-0.0620	-0.0519	-0.0596	-0.0496	-0.0657	-0.0555
	(0.0649)	(0.0651)	(0.0648)	(0.0650)	(0.0647)	(0.0650)
MiddleFiling	0.0309	0.0312	0.0300	0.0304	0.0323	0.0326
	(0.0523)	(0.0525)	(0.0523)	(0.0525)	(0.0522)	(0.0524)
AvgUnderpricing	-0.00321	-0.00171	-0.00295	-0.00152	-0.00355	-0.00217
	(0.0122)	(0.0122)	(0.0122)	(0.0122)	(0.0122)	(0.0122)
BMktshare	1.009	1.346	1.318	1.648	1.445	1.725
	(1.330)	(1.332)	(1.314)	(1.316)	(1.285)	(1.288)
MktReturn	0.222	0.321	0.272	0.372	0.220	0.320
	(0.639)	(0.642)	(0.639)	(0.641)	(0.638)	(0.641)
YearDummy	YES	YES	YES	YES	YES	YES
IndustryDummy	YES	NO	YES	NO	YES	NO
Constant	0.0634	0.0742	-0.0794	-0.0513	-0.0573	-0.0355
	(0.636)	(0.602)	(0.618)	(0.585)	(0.645)	(0.610)
Sample Size	1834	1834	1834	1834	1834	1834
R^2	0.085	0.074	0.083	0.073	0.085	0.075
adj. R^2	0.065	0.057	0.064	0.056	0.066	0.058

1. Dependent variable is adj_HPR1, the industry-adjusted holding period return in the first post-IPO year.

2. Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 11: Specification Results: Effect of Firm Centralities on HPR3

	(1)	(2)	(3)	(4)	(5)	(6)
Fdegree	0.688	0.697	0.688	0.695		
	(0.455)	(0.455)	(0.454)	(0.454)		
Feigenvector	-1447.4*	-1472.1*	-1472.6*	-1509.8**		
	(780.6)	(780.7)	(768.8)	(768.9)		
Fcloseness					-706.8	-763.2
					(704.6)	(704.2)
Fbetweenness	-195.2	-309.2			-191.4	-275.5
	(1139.2)	(1137.3)			(1156.9)	(1155.9)
Udegree	0.0580	0.0476	-0.0128	-0.00319		
	(0.324)	(0.324)	(0.298)	(0.298)		
Ueigenvector	-61.14	-56.92	-57.51	-54.13		
	(105.0)	(105.1)	(104.8)	(104.8)		
Ucloseness					123.8	118.0
					(550.8)	(551.4)
Ubetweenness	-4.429	-3.131			-7.072	-5.778
	(8.023)	(8.017)			(7.463)	(7.452)
LnOffersize	-0.0194	-0.00670	-0.0174	-0.00721	-0.0322	-0.0197
	(0.107)	(0.107)	(0.106)	(0.106)	(0.106)	(0.106)
HiTechDummy	0.267***	0.238***	0.270***	0.241***	0.265***	0.240***
	(0.0840)	(0.0769)	(0.0838)	(0.0767)	(0.0836)	(0.0765)
LnPrimaryShares	0.0189	0.00297	0.0153	0.000440	0.0356	0.0198
	(0.146)	(0.146)	(0.146)	(0.146)	(0.145)	(0.145)
FilingWidth20Dummy	0.341***	0.350***	0.340***	0.349***	0.338***	0.348***
	(0.116)	(0.116)	(0.116)	(0.116)	(0.116)	(0.116)
MiddleFiling	-0.0895	-0.0885	-0.0890	-0.0885	-0.0891	-0.0880
	(0.102)	(0.102)	(0.102)	(0.102)	(0.102)	(0.102)
AvgUnderpricing	0.00799	0.00452	0.00758	0.00425	0.00772	0.00439
	(0.0230)	(0.0230)	(0.0230)	(0.0230)	(0.0231)	(0.0231)
BMktshare	1.927	2.294	1.700	2.132	1.551	1.948
	(2.357)	(2.355)	(2.320)	(2.316)	(2.280)	(2.277)
MktReturn	-0.769	-0.650	-0.785	-0.672	-0.789	-0.670
	(1.142)	(1.142)	(1.140)	(1.140)	(1.142)	(1.142)
YearDummy	YES	YES	YES	YES	YES	YES
IndustryDummy	YES	NO	YES	NO	YES	NO
Constant	0.115	0.236	0.140	0.288	0.0584	0.210
	(1.124)	(1.059)	(1.088)	(1.025)	(1.144)	(1.080)
Sample Size	1518	1518	1518	1518	1518	1518
R^2	0.084	0.080	0.084	0.079	0.082	0.078
adj. R^2	0.061	0.058	0.062	0.060	0.060	0.058

1. Dependent variable is adj_HPR3, the industry-adjusted holding period return in the third post-IPO year.
2. Standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

3 PEER EFFECTS IN LATENT NETWORKS

Peer effects are widely studied for various aspects of network outcomes(Calvó-Armengol, Patacchini and Zenou (2009), Goldsmith-Pinkham and Imbens (2013)). Previous literature on identification of peer effects parameters mainly require the full knowledge of adjacency matrices that describe the network structures. However, in applied work, it is very likely to have a network with mismeasured links. For example, widely used self-reported survey data can contain intentional errors. Survey limitation on the number of links that one can report can lead to bounded degree of an observed network. In some network studies, mismeasured links arise because researchers assume the population can be partitioned into several groups and everyone in the same group are connected. This paper theoretically and empirically studies identification and estimation of peer effects in networks where mismeasured links are considered.

There are several undesired consequences if an observed network under interest has mismeasured links. Firstly, the peer effects estimated based on the observed network are biased if mismeasured links are not taken into consideration. Because mismeasured links can result in misclassification of link status from either 1 to 0 or 0 to 1 in the associated adjacency matrix, then both the mean of peer characteristics or the mean of peer outcomes may have measurement errors. Therefore, the peer effect estimated is biased. Secondly, if one studies information diffusion along links, then mismeasured links can lead to wrong diffusion paths. Because of misclassification on link status, the observed neighbors of one node is different from the latent neighbors of that node, information that passes through the node will end in different diffusion paths from the true paths. Moreover, if one is interested in exploring network formation, the mismeasured links can result in biased preference parameters. In the dynamic network formation study, individuals make decisions on forming a link or deleting a link based on comparing the expected utility between the current stage and the next stage. If the networks we observe over time involve in mismeasured links, then the estimates of the preference parameters in the utility function would be biased.

The objective of this paper is to both theoretically and empirically study the identification and estimation strategy of peer effects in networks, given that the econometricians may not be able to observe the latent network because of mismeasured links. In the existing literature, identification of peer effects in a given network has been studied (Blume et al. (2015), Bramoullé, Djebbari and Fortin (2009)). Nevertheless, how to identify and estimate the peer effects in networks with mismeasured links remains unclear.

To identify peer effects in networks with mismeasured links, we take three steps. In the first step, we investigate the identification strategy in a generalized linear-in-mean social interaction model. We use two adjacency matrices that summarize the link structures in peer-effects network and contextual-effects network, separately. In the peer-effects network, an individual's outcome is affected by her linked individual's outcome, while in the contextual-effects network, an individual's outcome is affected by her linked individual's characteristics. The peer-effects network and the contextual-effects network are two different network structures based on a same group of individuals. The outcome variable is linearly determined by the mean of neighbors' characteristics that in the contextual-effects network and the mean of peer outcomes that in the peer-effects network. Because of the linear-in-mean model, the misclassification of linkage status in the adjacency matrices are transformed to measurement errors in the mean outcomes and the mean characteristics. We first show that an weighted average of characteristics of neighbors in an individual's peers' contextual-effects network can be used as an instrument variable for the mean of that individual's peer outcomes (Bramoullé, Djebbari and Fortin (2009), Blume et al. (2015)). For simplicity, we call the instrument variable for the mean of peer outcomes as instrumental mean characteristics. The social effects can be identified if the observed network is assumed to be the true network of interest.

Then to identify peer effects in networks with mismeasured links, we first apply the method in Hu (2008). We use repeated measurements on two conditional distribution matrices: one is the joint distribution matrix of observed mean characteristics/instrumental mean characteristics and outcomes conditional on covariates; the other is the distribution matrix of observed mean characteristics/instrumental mean characteristics conditional on covariates. Then we use an eigenvalue-eigenvector decomposition method based on above distribution matrices to obtain conditional densities of latent mean characteristics/instrumental mean characteristics conditional on observed mean characteristics/instrumental mean characteristics, separately. We use Gumbel copula parameter, along with the two conditional densities, to get the joint distribution of latent mean characteristics and latent instrumental mean characteristics. The intuition is that repeated measurements on mean characteristics/instrumental mean characteristics serve as instrument variables, so that we can identify conditional densities of latent means conditional on observed means. Finally, we adopt the conditional densities along with latent moment conditions to derive an observed conditional model. Then we use a GMM method to estimate peer effects parameters. In summary, the intuition is that with the help of conditional densities of latent instrumental mean characteristics, we can identify peer effects in a GMM model through an observed conditional moment model, instead of knowing the latent network structures.

In the second part, we develop a three-stage semiparametric estimator following the identification procedure. In the first stage, we use repeated measurements to estimate the conditional densities of mean characteristics and instrumental mean characteristics, separately(Hu (2008)). In the second stage, we propose a semiparametric pseudo-likelihood method to estimate a Gumbel copula parameter (Genest, Ghoudi and Rivest (1995)). We use this copula parameter, along with the distribution matrices in the first stage, to get the joint distribution of latent mean characteristics and latent instrumental mean characteristics conditional on covariates. In the third stage, we can estimate the social effects parameters by a GMM method.

In the third part, we carry out an empirical study by constructing firm networks then using a structural form analysis. In our empirical study, to construct the peer-effects network and contextual-effects network, we use data from the Customer-supplier data and the merged Center for Research in Security Prices (CRSP)-Compustat database. The firms' peer-effects network is constructed by defining the linkages between two firms based on if they are customer and supplier links. It has been seen in the finance literature that the choices made by the customers do affect the choices made by the suppliers, and vice versa. The firms' contextual-effects network is constructed by defining the linkages between two firms based on if they have the same industry identification code. It is quite clear that the firms that are in the same industry have similar characteristics and we think these similarities may have impact on firms' choices. As required by our identification strategy, we need at least two repeated measurements of both the peer-effects network and the contextual-effects network. For the peer-effects network, we use the observed customer-supplier network in 2001 as the first measurement, while that in 2002 as the second measurement. As for the contextual-effects network, we use the SIC 3-digit industry classification for constructing the first measurement of adjacency matrix, while the GIC industry classification for the second measurement. Our dependent variable is the leverage ratio chosen by each firms, a measure of firm's corporate financial policy. Our control variables include measures for firm capital structure and measures for firm profitability. Following Leary and Roberts (2014), we use the return shock as the exogenous peer firm characteristic that serves as the key independent variable. We begin with a known capital structure determinant, stock returns, for the time period Jan.1991 to Dec.2000, then we use a traditional asset pricing model to extract the idiosyncratic variation, along with the variation among peers in stock returns. We obtain the residual from this model and define this residual as return shock and use it to obtain the exogenous variation in peer firms' characteristics in year 2001. Such exogenous variation in year 2001 is used as our key independent variable. At last we adopt the three-stage estimation method to obtain the estimates of structural parameters on peer effect and contextual effect.

Empirical findings indicate that the contextual effect and the peer effect among firms are statistically significant. The contextual effect is positive significant with a scale around 0.3. This indicates that if an industry experiences an increasing return shock, it is very likely that the firms in the same industry to choose higher leverage ratio. A positive return shock usually indicates a prosperity in the industry, the firms in the same industry choose to increase leverages in order to invest and gain more from the booming stock market. The peer effect is also found to be positive and significant, with a scale around 0.8. This implies that if a firm's customer or supplier choose to increase its leverage ratio, it is very likely that the firm would also choose to increase its leverage ratio. Indicated by the bargaining power theory, when firms raise their leverages, their suppliers of customers will also raise their own leverages in response to increase their bargaining power (Hennessy and Livdan (2009), Chu (2012)). Furthermore, the peer effect becomes insignificant and the magnitude decreases if we ignore the possibilities of mismeasured network links.

This paper contributes to the existing literature of peer effects in network in several aspects. First, this paper is one of the first papers studying the identification and estimation of peer effects in networks with mismeasured links. In the existing literature, the adjacency matrix of network structures is required to be observed in the sample when identifying and estimating the network effects parameters (Bramoullé, Djebbari and Fortin (2009), Blume et al. (2015)). There are few studies on identification of network effects parameters or preference parameters under certain restrictions on the observed links. In De Paula, Richards-Shubik and Tamer (2018), they study the identification of preference parameters in networks with bounded degree. In Chandrasekhar and Lewis (2011), they investigate the identification of social networks with use of sampled network data. The problem of mismeasured links arises because in applied work, researchers generally collect a sub-sample of a network. In their paper, the links between observed individuals in the sub-sample are assumed to be true links. However, in this paper, we analyze peer effects under the case where all the links in a network could be mismeasured with an assumption that the supports of latent mean characteristics/instrumental mean characteristics are the same as the the supports of observed mean characteristics/instrumental mean characteristics. In Blume et al. (2015), they mention the idea on how to deal with unknown social matrix. Our paper provides a novel sufficient conditions of the identification strategy on peer effect parameters that are different from the above literature.

Second, we contribute to the empirical literature by applying our method to analyze peer effects of firms on their financial policies. In the existing literature, peer effects are often studied under a network structure where firm linkages are often constructed based on if two firms are in the same industry (Leary and Roberts (2014)). However, it is very unlikely to have every firm in the same industry be connected and affect each other's actions. Besides, for different kinds of industry classification code (e.g. SIC industry code, GIC code, Standard Poor industry classification, etc.), the classifications on firms are different, as a result the associated firm networks are different. Other applied work on firm networks use data that to some extent describe the relationship between two firms (e.g. customer-supplier data, connection data of members in the boards). However, for a static network study, such data may be just a instantaneous observation on the real network. There can be firms that are linked but unobserved and also firms that are only interacted for a very short period but still under observed. Hence in the study of firm networks, the problem of mismeasured links requires a lot attention. Besides, many existing econometric studies on estimation of network effects use social network data, such as the Add Health data, the 75 villages in India data, and etc (as in e.g., Bramoullé, Djebbari and Fortin (2009), Calvó-Armengol, Patacchini and Zenou (2009),Goldsmith-Pinkham and Imbens (2013), **?**, etc.) By using the customer-supplier data and the CRSP data, we study peer effects among firms on their corporate financial policy(Leary and Roberts (2014)) and provide new evidence on peer effects in economic networks.

The rest of the paper is organized as follows. Section 2 introduces the linear social interaction model we use. Section 3 and 4 shows the identification strategy of peer effects in networks without and with latent links, respectively. Section 5 describes the three-stage semi-parametric estimation method. Section 6 studies the asymptotic properties of the three-stage semi-parametric estimators. Section 7 applies the estimation method on firm networks and empirically studies the peer effects on firms' financial policies. Section 8 discusses conclusions and potential future works.

3.1 The Model

We consider a generalized linear-in-mean social interaction model where individual outcome Y_i is determined by

$$Y_i = \alpha + \beta \sum_j A^*_{ij} Y_j + \gamma X^c_i + \delta \sum_j C^*_{ij} X_j + \varepsilon_i. \tag{3.1}$$

In this model, X^c_i is a $1 \times k$ vector controlling for individual i's exogenous characteristics and $X^c \equiv (X^c_1, X^c_2, ..., X^c_n)^T$ is a $n \times k$ matrix of characteristics. Let A^* and C^* be two row-normalized symmetric $n \times n$ adjacency matrices that summarize the link status in *peer-effects* network and *contextual effects* network, with $A^*_{ij} = 1/n_i$ if i and j are linked and $A^*_{ij} = 0$ otherwise, where n_i is the number of individuals that are linked to i in the peer-effects network. C^*_{ij} is defined similarly to A^*_{ij}. Let $A^*_{ii} = C^*_{ii} = 0$ by convention in the literature that individuals cannot link to themselves. Similar to Blume et al. (2015), A^* can be different from C^*. We assume that no individuals are isolated in the network so that the row-normalization of A^* and C^* are well-defined. Note that we use the superscript "$*$" to emphasize the point that A^* and C^* are latent to econometrician because of mismeasured links.

Based on the definitions of A^*_{ij} and C^*_{ij}, the covariate $\sum_j A^*_{ij} Y_j$ represents the average outcomes of individuals in i's peer-effects network. For simplicity, we call it as latent mean outcomes. The term $\sum_j C^*_{ij} X_j$ denotes the average characteristics of individuals in i's contextual-effects network, which we assume to be a discrete scalar and exogenous. Similarly, we call it as latent mean characteristics. ε_i is the i.i.d error term. The structural parameters of interest include the intercept $\alpha \in \mathbb{R}$, the endogenous peer effect $\beta \in \mathbb{R}$, the vector of individual effects $\gamma \in \mathbb{R}^k$, and the contextual effects $\delta \in \mathbb{R}$. Following the literature, we require that $|\beta| < 1$.

This paper is to identify and estimate the social effects parameters $\theta \equiv (\alpha, \beta, \gamma^T, \delta)^T$, given that the econometrician may not be able to observe the true network under interest because of mismeasured links.

3.1.1 Identification without Latent Networks

In this section, we identify θ by assuming that A^* and C^* are known to analyst a priori. We show that identification is established by imposing a simple condition on matrices A^* and C^*.

We can write the structural model (3.1) in matrix notation, which is

$$Y = \alpha\iota + \beta A^* Y + \gamma X^c + \delta C^* X + \varepsilon, \tag{3.2}$$

where Y is an $n \times 1$ vector of actions for all individuals in the network, ι is an $n \times 1$ vector of ones. Since $|\beta| < 1$, $(I - \beta A^*)$ is invertible[1]. We can write:

$$Y = \alpha(I-\beta A^*)^{-1}\iota + \gamma(I-\beta A^*)^{-1}X^c + \delta(I-\beta A^*)^{-1}C^*X + (I-\beta A^*)^{-1}\varepsilon.$$

By Neumann expansion, $(I-\beta A^*)^{-1} = \sum_{s=0}^{\infty}\beta^s A^{*s}$. Then by assuming no isolated individuals in the latent peer-effects network A^*[2], we have

$$Y = \frac{\alpha}{1-\beta}\iota + \gamma\sum_{s=0}^{\infty}\beta^s A^{*s} X^c + \delta\sum_{s=0}^{\infty}\beta^s A^{*s} C^* X + \sum_{s=0}^{\infty}\beta^s A^{*s}\varepsilon, \tag{3.3}$$

which can be viewed as the reduced from equation for (3.2).

Assumption 3.1. $\mathbb{E}(\varepsilon_i | X^c, X) = 0$.

Assumption 3.1 means that individual characteristics X^c, X are strictly exogenous. Then latent mean outcomes conditional on X^c, X can be written as:

$$\mathbb{E}(A^*Y | X^c, X) = \frac{\alpha}{1-\beta}\iota + \gamma\sum_{s=0}^{\infty}\beta^s A^{*s+1} X^c + \delta\sum_{s=0}^{\infty}\beta^s A^{*s+1} C^* X. \tag{3.4}$$

Equation (3.4) implies that variables $(A^*C^*X, A^{*2}C^*X, \dots)$ can be used as instrumental variables

[1] See Case (1991), footnote 5.

[2] This implies that $A^*\iota = \iota$ and $\alpha(I-\beta A^*)^{-1}\iota = \alpha/(1-\beta)\iota$.

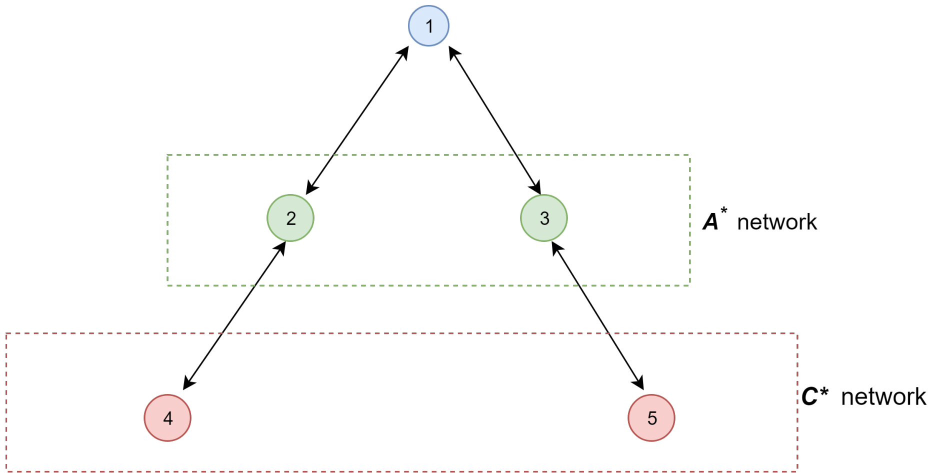


Figure 5: Illustration of IV

for A^*Y, provided that $\delta \neq 0$. These IVs have socioeconomic interpretation. For example, A^*C^*X represents an $n \times 1$ vector of weighted averages of characteristics for members in peers' contextual-effects network. Figure 5 illustrates the intuition for adopting A^*C^*X as IV. In this figure, individual 1 is linked to 2 and 3 in the peer-effects network, and 2 and 3 are linked to 4 and 5 in the contextual-effects network, respectively. We are interested in peer effects of 2 and 3's outcomes on 1's outcome. Due to the network structure, the effects of 4 and 5's characteristics on 1's outcome are solely because of their effects on 2 and 3's outcomes, which forms a valid exclusion restriction.

Proposition 3.1. *Suppose that $\delta \neq 0$. If the matrices I, A^*, C^* and A^*C^* are linearly independent, the social effects θ are identified.*

Sufficient conditions that depend only on the shape of networks for linear independence of I, A^*, C^*, A^*C^* are given by Lemma 3.1 below and are easily satisfied.

Lemma 3.1. *If there exists two distinct individuals i and j such that (i) $\sum_k A^*_{ik}C^*_{ki} \neq \sum_k A^*_{jk}C^*_{kj}$ or (ii) they are linked by a sequence of edges, some in the peer-effects network and some in the contextual-effects network, and who are not linked by a path in either the peer-effects network or the contextual-effects network alone, then I, A^*, C^* and A^*C^* are linearly independent.*

In order to understand this lemma, suppose that social interaction are confined to firms in some industries. The term $\sum_k A^*_{ik} C^*_{ki}$ represents the contextual effect of firms that are linked to firm i in its peer-effects network, and condition (i) requires this effect to be different for two distinct firms. This condition can be easily satisfied if contextual effects are not distributed uniformly across all firms. Condition (ii), on the other hand, is satisfied except possibly when the peer-effects network and contextual-effects network are the same, i.e., $A^* = C^*$.

Define $Z^*_i = \left(\sum_{j=1}^n C^*_{ij} X_j, \sum_{j=1}^n \sum_{k=1}^n A^*_{ij} C^*_{jk} X_k \right)$, same in matrix form $Z^* = (C^* X, A^* C^* X)$.

ASSUMPTION 3.2. *(i) $A^{*s} X^c$ and $A^{*t} C^* X^c$ are mean independent of $C^* X$ and $A^* C^* X$, where $s = 1, 2, \ldots$, $t = 2, 3, \ldots$. (ii) $E\left[\sum_{s=1}^\infty \beta^s E(A^{*s} X^c) + \sum_{s=2}^\infty \beta^t E(A^{*t} C^* X^c)\right] = 0$*

In Assumption 3.2(i), consider the case where $s = 1$ and $t = 2$, that assuming $A^* X^c$ and $A^* C^* X^c$ are mean independent of $C^* X$ and $A^* C^* X$. $A^* X^c$ is the mean of covariates in the peer-effects network, while $A^* C^* X^c$ is the mean of instrumental covariates. Then Assumption 3.2(i) implies that, in the latent networks, mean covariates and instrumental mean covariates are mean independent of mean characteristics and instrumental mean characteristics. Because of mean independence, $E(Y|X^c, Z^*)$ can be written as

$$E(Y|X^c, Z^*) = \frac{\alpha}{1-\beta}\iota + \gamma X^c + \delta C^* X + \delta\beta A^* C^* X + E\left[\sum_{s=1}^\infty \beta^s E(A^{*s} X^c) + \sum_{t=2}^\infty \beta^t E(A^{*t} C^* X^c)\right]$$

Based on zero mean condition 3.2(ii), the reduced form equation (3.3), and discussions above, we have the following parametric conditional moment model:

$$\mathbb{E}(Y_i|X^c_i, Z^*_i) = \frac{\alpha}{1-\beta} + \gamma X^c_i + \delta \sum_{j=1}^n C^*_{ij} X_j + \delta\beta \sum_{j=1}^n \sum_{k=1}^n A^*_{ij} C^*_{jk} X_k \equiv m(X^c_i, Z^*_i; \theta). \tag{3.5}$$

Therefore, we can identify the social parameters θ by the conditional moment condition

$$\mathbb{E}\left[Y_i - m(X^c_i, Z^*_i; \theta) | X^c_i, Z^*_i\right] = 0. \tag{3.6}$$

3.1.2 Identification with Latent Networks

In this section, we discuss identification of social parameters θ when A^* and C^* are unobserved because of mismeasured links. Then the moment condition (3.6) becomes infeasible. In data, we can only observe contaminated measures A and C for A^* and C^*, respectively. Let Z_i denote the (observed) instrument variables with A^*_{ij} replaced by A_{ij} and C^*_{ij} replaced by C_{ij}

Let $f_{Z^*|Z,X^c}$ denote the conditional density of latent variable given the contaminated measure Z and the covariate X^c. Define $Z_i = \left(\sum_{j=1}^n C_{ij}X_j, \sum_{j=1}^n \sum_{k=1}^n A_{ij}C_{jk}X_k\right)$, where A_{ij} is the entry in the i-th row and j-th column of matrix A (and similarly for C_{ij}). If we can identify $f_{Z^*|Z,X^c}$, then instead of the latent model (3.5), we can identify the social parameters θ through the observed conditional moment model

$$\mathbb{E}(Y_i|X_i^c, Z_i) = \int m(X_i^c, Z_i^*; \theta) f_{Z^*|Z,X^c} dZ^*. \tag{3.7}$$

The above observed conditional moment model (3.7) holds under Assumption 3.4(ii), proof can be found in Appendix B. It implies, by the law of iterated expectation, the following unconditional moment condition:

$$\mathbb{E}\left[\left(Y_i - \int m(X_i^c, Z_i^*; \theta) f_{Z^*|Z,X^c} dZ^*\right) S_i\right] \equiv \mathbb{E}\left[g(Y_i, S_i, \theta)\right] = 0, \tag{3.8}$$

where $S_i = (1, X_i^{cT}, Z_i^T)^T$. Therefore, the objective of this section is to identify the conditional density $f_{Z^*|Z,X^c}$.

Denote by $\mathcal{Z}^*$ and $\mathcal{Z}$ the support of Z^* and Z. Let $Z_1^* \equiv \sum_{j=1}^n C^*_{ij}X_j$ and $Z_2^* \equiv \sum_{j=1}^n \sum_{k=1}^n A^*_{ij}C^*_{jk}X_k$ (and similarly for Z_1 and Z_2). By Bayes' Theorem,

$$f_{Z^*|Z,X^c} = \frac{f_{Z|Z^*,X^c} f_{Z^*|X^c} f_{X^c}}{f_{Z,X^c}}. \tag{3.9}$$

Therefore, in order to identify $f_{Z^*|Z,X^c}$ we can try to identify $f_{Z|Z^*,X^c}$ and $f_{Z^*|X^c}$.

ASSUMPTION 3.3. *(i)* $f_{Z|Z^*,X^c} = f_{Z_1|Z^*,X^c} f_{Z_2|Z^*,X^c}$ *and (ii)* $f_{Z_m|Z^*,X^c} = f_{Z_m|Z_m^*,X^c}$ *for* $m = 1,2$*.*

Assumption 3.3 (i) means that Z_1 and Z_2 are independent with each other when conditioning on the latent variables Z^* and the covariate X^c. A sufficient condition for condition (i) is the independence of measurement errors in Z_1 and Z_2. For example, suppose $Z_1^* = (X_1 + X_2 + X_3)/3$, $Z_1 = (X_1 + X_2 + X_4)/3$, $Z_2^* = (X_5 + X_6)/2$ and $Z_2 = (X_6 + X_7)/2$. Then the measurement errors in Z_1 and Z_2 are $(X_4 - X_3)/3$ and $(X_7 - X_5)/2$, respectively. The measurement errors are independent with each other because $\{X_i\}_{i \in N}$ is independent across different individuals.

Condition (ii) requires Z_l to be independent with Z_m^* when conditioning on Z_l^* and X^c for $l, m = 1,2$, $l \neq m$. One of the sufficient conditions for (ii) is that the measurement error in Z_l is independent with Z_m^*. Consider the example above, the measurement error in Z_1 is $(X4 - X3)/3$, which is independent of $Z_2^* = (X_5 + X_6)/2$.

By Assumption 3.3,

$$f_{Z|Z^*,X^c} = f_{Z_1|Z_1^*,Z_2^*,X^c} f_{Z_2|Z_1^*,Z_2^*,X^c} = f_{Z_1|Z_1^*,X^c} f_{Z_2|Z_2^*,X^c}.$$

Consequently, in order to identify $f_{Z|Z^*,X^c}$ we can try to identify the conditional densities $f_{Z_1|Z_1^*,X^c}$ and $f_{Z_2|Z_2^*,X^c}$. The identification requires another measurement of the peer-effects and contextual-effects network, which we denote as A' and C'.

Then we have $Z_1' \equiv \sum_{j=1}^n C_{ij}' X_j$ and $Z_2' \equiv \sum_{j=1}^n \sum_{k=1}^n A_{ij}' C_{jk}' X_k$.We impose the following additional assumptions to in order to apply the results in Hu (2008).

ASSUMPTION 3.4. *For* $m = 1,2$*, (i)* $f_{Z_m|Z_m',Z_m^*,X^c} = f_{Z_m|Z_m^*,X^c}$ *; (ii)* $f_{Y|Z_m,Z_m',Z_m^*,X^c} = f_{Y|Z_m^*,X^c}$

Assumption 3.4 (i) requires the two measurements to be independent with each other when conditioning on the latent variable $Z_m^*, m = 1,2$ and the covariate X^c. Equivalently, it requires the measurement error in Z_1 is independent of the measurement error in Z_1'. For an arbitrary individual i, this means its mismeasured links in different observed networks should be different. Let us still use the example discussed before: $Z_1^* = (X_1 + X_2 + X_3)/3$ and $Z_1 = (X_1 + X_2 + X_4)/3$. If

$Z_1' = (X_2 + X_3 + X_5)/3$, then the measurement errors in Z_1 and Z_1' are $(X_4 - X_3)/3$ and $(X_5 - X_1)/3$, which are obviously independent with each other. Condition (ii) means the two measurements are independent with the outcome variable Y when conditioning on the latent variable and the covariate. Assumption 3.4 is commonly imposed in the measurement error literature, see, e.g., Hu (2008), Hu and Schennach (2008) and Hu (2017).

ASSUMPTION 3.5. $|\mathcal{Z}_m| = |\mathcal{Z}_m'| = |\mathcal{Z}_m^*| = K_m$, $m = 1, 2$.

Assumption 3.5 is similar to Hu (2017) and requires the dimension of the support of two measurements to be equal to that of the support of Z_m^*. Otherwise the measurements will not have enough information to identify the distribution of the latent variables. This assumption can be relaxed to allow $|\mathcal{Z}_m|$ and $|\mathcal{Z}_m'|$ to be larger than $|\mathcal{Z}_m^*|$. Assumptions 3.4 implies that

$$\begin{aligned} &f_{Z_m,Z_m',Y|X^c}(z_m, z_m', y|x^c) = \\ &\sum_{z_m^* \in \mathcal{Z}^*} f_{Z_m|Z_m^*,X^c}(z_m|z_m^*, x^c) f_{Z_m'|Z_m^*,X^c}(z_m'|z_m^*, x^c) f_{Y|Z_m^*,X^c}(y|z_m^*, x^c) f_{Z_m^*|X^c}(z_m^*|x^c). \end{aligned} \tag{3.10}$$

For each $Y = y$ and $X^c = x^c$, we define

$$M_{Z_m,Z_m',Y|X^c} = \left[f_{Z_m,Z_m',Y|X^c}(z_{mi}, z_{mj}', y|x^c) \right]_{i=1,2,\dots,K_m; j=1,2,\dots,K_m},$$

$$D_{Y|Z_m^*,X^c} = \operatorname{diag}\left\{ f_{Y|Z_m^*,X^c}(y|z_{m1}^*, x^c), f_{Y|Z_m^*,X^c}(y|z_{m2}^*, x^c), \dots, f_{Y|Z_m^*,X^c}(y|z_{mK_m}^*, x^c) \right\},$$

$$M_{Z_m|Z_m^*,X^c} = \left[f_{Z_m|Z_m^*,X^c}(z_{ml}|z_{mk}^*, x^c) \right]_{l=1,2,\dots,K_m; k=1,2,\dots,K_m},$$

$$D_{Z_m^*|X^c} = \operatorname{diag}\left\{ f_{Z_m^*|X^c}(z_{m1}^*|x^c), f_{Z_m^*|X^c}(z_{m2}^*, x^c), \dots, f_{Z_m^*|X^c}(z_{mK_m}^*|x^c) \right\}.$$

Then (3.10) is equivalent to

$$M_{Z_m,Z'_m,Y|X^c} = M_{Z_m|Z^*_m,X^c} D_{Y|Z^*_m,X^c} D_{Z^*_m|X^c} M^T_{Z'_m|Z^*_m,X^c}, m = 1,2. \tag{3.11}$$

ASSUMPTION 3.6. *$M_{Z_m|Z^*_m,X^c}$ and $M_{Z'_m|Z^*_m,X^c}$ has rank K_m for $m = 1,2$.*

Assumption 3.6 has been used in Newey and Powell (2003), Mahajan (2006) and Hu (2008). In the case when the measurements Z_m and Z'_m take fewer values than Z^*_m, Assumption 3.6 may fail to hold. In Appendix 5, we give an example of how easy this assumption can be satisfied. Define

$$M_{Z_m,Z'_m|X^c} = \left[f_{Z_m,Z'_m|X^c}(z_{mi}, z'_{mj}|x^c)\right]_{i=1,2,\ldots,K_m; j=1,2,\ldots,K_m}.$$

Similar to (3.11), we have

$$M_{Z_m,Z'_m|X^c} = M_{Z_m|Z^*_m,X^c} D_{Z^*_m|X^c} M^T_{Z'_m|Z^*_m,X^c}. \tag{3.12}$$

Equation (3.12) implies that the rank of the observed matrix $M_{Z_m,Z'_m|X^c}$ equals to K_m. Therefore, Assumption 3.6 can be tested.

Lemma 3.2. *Under Assumption 3.5 and 3.6, the rank of matrix $M_{Z_m,Z'_m|X^c}$ is K_m for $m = 1,2$.*

Assumption 3.6 together with (3.11) and (3.12) implies that

$$M_{Z_m,Z'_m,Y|X^c} M^{-1}_{Z_m,Z'_m|X^c} = M_{Z_m|Z^*_m,X^c} D_{Y|Z^*_m,X^c} M^{-1}_{Z_m|Z^*_m,X^c}. \tag{3.13}$$

This equation implies that observed matrix on the left hand side of (3.13) has an eigenvalue-eigenvector decomposition. Then $f_{Z_m|Z^*_m,X^c}$ can be identified as the eigenvector of the observed matrix up to permutation of its rows. In order to guarantee the identification is unique, we need to have another assumption:

ASSUMPTION 3.7. *For $m = 1,2$ and $z_m^* \neq \widetilde{z}_m^*$, (i) $f_{Y|Z_m^*,X^c}(y|z_m^*,x^c) \neq f_{Y|Z_m^*,X^c}(y|\widetilde{z}_m^*,x^c)$; (ii) $f_{Z_m|Z_m^*,X^c}(z_m^*|z_m^*,x^c) > f_{Z_m|Z_m^*,X^c}(\widetilde{z_m}^*|z_m^*,x^c)$.*

Assumption 3.7 (i) ensures that the eigenvalues $\{f_{Y|Z_m^*,X^c}(z_m''|z_{mk}^*,x^c)\}_{k=1,2,...,K_m}$ are distinctive and rules out the case of duplicate eigenvalues. Condition (ii) implies that z_m^* is the mode of distribution $f_{Z_m|Z_m^*,X^c}(y|z_m^*,x^c)$ and hence the ordering of the eigenvectors. These two conditions together guarantee that the decomposition in (3.13) is unique.

The conditional marginal distribution of Z_m can be represented as $f_{Z_m|X^c} = \sum_{Z_m^*} f_{Z_m|Z_m^*,X^c} f_{Z_m^*|X^c}$. The matrix representation is

$$M_{Z_m|X^c} = M_{Z_m|Z_m^*,X^c} M_{Z_m^*|X^c},$$

where

$$M_{Z_m|X^c} = \left[f_{Z_m|X^c}(z_{ml}|x^c)\right]^T_{l=1,2,...,K_m}$$

and

$$M_{Z_m^*|X^c} = \left[f_{Z_m^*|X^c}(z_{ml}^*|x^c)\right]^T_{l=1,2,...,K_m}.$$

Therefore, by Assumption 3.6 the conditional density of Z_m^* can be identified as

$$M_{Z_m^*|X^c} = M^{-1}_{Z_m|Z_m^*,X^c} M_{Z_m|X^c},\ m = 1,2. \tag{3.14}$$

Proposition 3.2. *Under Assumptions 3.4-3.7, the conditional densities $f_{Z_m|Z_m^*,X^c}$ and $f_{Z_m^*|X^c}$ are nonparametrically identified for $m = 1,2$.*

The next step is to identify the (conditional) joint density $f_{Z^*|X^c}$. Since Z_1^* and Z_2^* can be dependent with each other[3], we will apply Sklar's Theorem to identify it. Let C_σ denote the parametric conditional copula function with parameter σ,

ASSUMPTION 3.8. *For conditional copula functions $C_\sigma\left(F_{Z_1^*|X^c}, F_{Z_2^*|X^c}\right)$ and $C_{\widetilde{\sigma}}\left(F_{Z_1|X^c}, F_{Z_2|X^c}\right)$,*

[3] Z_1^* and Z_2^* may contain same individual's characteristics.

$\sigma = \widetilde{\sigma}$.

Assumption 3.8 requires the parameter of the conditional copula function for Z_1^* and Z_2^* to be the same as that of Z_1 and Z_2. This assumption implies that the dependent structure Z_1^* and Z_2^* will not be contaminated by the measurement errors. Hence, the copula parameter σ can be identified by the conditional distributions of Z_1 and Z_2. Furthermore, the conditional distribution functions of Z_m^* can be identified by $f_{Z_m^*|X^c}$. Consequently, by Sklar's Theorem we have

$$f_{Z^*|X^c} = c_\sigma\left(F_{Z_1^*|X^c}, F_{Z_2^*|X^c}\right) f_{Z_1^*|X^c} f_{Z_2^*|X^c},$$

where c_σ is the density of C_σ. Finally, because the densities f_{X^c} and f_{Z,X^c} can be directly identified from data, the (joint) conditional density $f_{Z^*|Z,X^c}$ can be identified by (3.9).

Theorem 3.1. *Under Assumptions 3.3-3.8, the conditional density $f_{Z^*|Z,X^c}$ is semiparametrically identified.*

3.2 Estimation

Since the identification is constructive, we propose a three-stage semiparametric estimator following the identification procedure step by step. In the first stage, we estimate the conditional densities $f_{Z_m|Z_m^*,X^c}$ and $f_{Z_m^*|X^c}$ by nonparametric method. In the second stage, the copula function C is estimated via the pseudo likelihood method similar to Genest, Ghoudi and Rivest (1995), and we obtain the conditional joint density of Z_1^* and Z_2^*. Finally, in the third stage we estimate the social parameters θ by GMM.

3.2.1 Nonparametric Estimation of the Conditional Density

We estimate the joint distributions of Z_m, Z'_m, Y and X^c using a simple frequency estimator ($m = 1,2$),

$$\widehat{f}_{Z_m,Z'_m,Y,X^c}(z_m,z'_m,y,x^c) = \frac{1}{n}\sum_{i=1}^{n} \mathbb{1}(Z_{mi} = z_m, Z'_{mi} = z'_m, Y_i = y, X^c_i = x^c),$$

where $\mathbb{1}(\cdot)$ is the indicator function. Similarly, we can estimate f_{Z_m,Z'_m,X^c} and f_{X^c} using the frequency estimator. Then, the conditional distribution matrices $M_{Z_m,Z'_m,Y|X^c}$ and $M_{Z_m,Z'_m|X^c}$ can be estimated by stacking the estimate of f_{Z_m,Z'_m,Y,X^c}, f_{Z_m,Z'_m,X^c} and f_{X^c} as follows:

$$\widehat{M}_{Z_m,Z'_m,Y|X^c} = \left[\frac{\widehat{f}_{Z_m,Z'_m,Y,X^c}(j,l,y,x^c)}{\widehat{f}_{X^c}(x^c)}\right]_{j,l},$$

and

$$\widehat{M}_{Z_m,Z'_m|X^c} = \left[\frac{\widehat{f}_{Z_m,Z'_m,X^c}(j,l,x^c)}{\widehat{f}_{X^c}(x^c)}\right]_{j,l}.$$

Then, following the identification procedure, the conditional density $f_{Z_m|Z^*_m,X^c}$ are estimated as

$$\widehat{f}_{Z_m|Z^*_m,X^c} = \psi\left(\widehat{M}_{Z_m,Z'_m,Y|X^c}\widehat{M}^{-1}_{Z_m,Z'_m|X^c}\right),$$

where $\psi(\cdot)$ is the eigenvector function. Note that we need to normalize the sum of each column to be 1. Furthermore, the marginal distribution Z^*_m conditional on X^c can be estimated by

$$\widehat{M}_{Z^*_m|X^c} = \widehat{M}^{-1}_{Z_m|Z^*_m,X^c}\widehat{M}_{Z_m|X^c},$$

where $\widehat{M}_{Z^*_m|X^c} = [\widehat{f}_{Z^*_m|X^c}(j|x^c)]_j$. Note that $\widehat{M}_{Z_m|X^c} = [\widehat{f}_{Z_m|X^c}(j|x^c)]_j$ is directly obtained from data by using the frequency estimator.

3.2.2 Semiparametric Estimation of the Copula Function

Following Genest, Ghoudi and Rivest (1995), we propose a semiparametric pesudo-likelihood method to estimate the copula function C. Let

$$\widehat{F}_{Z_m,X^c}(z_m,x^c) = \frac{1}{n}\sum_{i=1}^{n} \mathbb{1}(Z_{mi} \leq z_m, X_i^c \leq x^c)$$

and

$$\widehat{F}_{X^c}(x^c) = \frac{1}{n}\sum_{i=1}^{n} \mathbb{1}(X_i^c \leq x^c)$$

denote the empirical distribution functions for F_{Z_m,X^c} and F_{Z_m,X^c}, $m = 1,2$. Then, the conditional distribution function $F_{Z_m|X^c}$ can be estimated as

$$\widehat{F}_{Z_m|X^c}(z_m|x^c) = \frac{\widehat{F}_{Z_m,X^c}(z_m|x^c)}{\widehat{F}_{X^c}(x^c)}.$$

We parametrize the conditional copula function to be C_σ. Then, the copula parameter σ will be estimated by

$$\widehat{\sigma} = \underset{\sigma}{\operatorname{argmax}} \sum_{i=1}^{n} \log\left[c_\sigma(\widehat{F}_{Z_1|X^c}(Z_{1i}|X_i^c), \widehat{F}_{Z_2|X^c}(Z_{2i}|X_i^c))\right], \tag{3.15}$$

where c_σ is the density of C_σ. Then, the identification procedure implies that the density $f_{Z^*|X^c}$ can be estimated as

$$\widehat{f}_{Z^*|X^c} = c_{\widehat{\sigma}}\left[\widehat{F}_{Z_1^*|X^c}, \widehat{F}_{Z_2^*|X^c}\right]\widehat{f}_{Z_1^*|X^c}\widehat{f}_{Z_2^*|X^c}.$$

3.2.3 GMM Estimation of the Social Parameters

In the third stage, we estimate the social parameters θ by the GMM method. Specifically, let $\widehat{V}$ denote a $(k+3)\times(k+3)$ weighting matrix,

$$\hat{\theta} = \underset{\theta}{\operatorname{argmin}} \left(\frac{1}{n}\sum_{i=1}^{n}\widehat{g}(Y_i,S_i,\theta)\right)^T \widehat{V}\left(\frac{1}{n}\sum_{i=1}^{n}\widehat{g}(Y_i,S_i,\theta)\right), \tag{3.16}$$

where

$$\widehat{g}(Y_i,S_i,\theta)=\left(Y_i-\int m(X_i^c,Z_i^*;\theta)\widehat{f}_{Z^*|Z,X^c}(z^*|Z_i,X_i^c)dz^*\right)S_i,$$

and

$$\widehat{f}_{Z^*|Z,X^c}=\frac{\widehat{f}_{Z|Z^*,X^c}\widehat{f}_{Z^*|X^c}\widehat{f}_{X^c}}{\widehat{f}_{Z,X^c}}.$$

3.3 Asymptotic Properties

In this section, we establish the asymptotic properties of the three-stage semiparametric estimators. First, we show that the frequency estimator for the conditional density is consistent.

ASSUMPTION 3.9. *(i)* $\{X_i^c\}_{i\in N}$ *is i.i.d; (ii)* $\{W_{mi}\}_{i\in N}\equiv\{(Z_{mi},Z'_{mi},Y_i)\}_{i\in N}$ *is strictly stationary for* $m=1,2$*; (iii) For* $m=1,2$*,* $\lim_{n\to\infty}\text{Cov}[\mathbb{1}(Z_{mi}=z,X_i^c=x),\mathbb{1}(Z_{mj}=z,X_j^c=x)]=0$*,* $\lim_{n\to\infty}\text{Cov}[\mathbb{1}(Z_{mi}=z,Z'_{mi}=Z,X_i^c=x),\mathbb{1}(Z_{mj}=z,Z'_{mj}=z,X_j^c=x)]=0$ *and* $\lim_{n\to\infty}\text{Cov}[\mathbb{1}(W_{mi}=w,X_i^c=x),\mathbb{1}(W_{mj}=w,X_j^c=x)]=0$*.*

Assumption 3.9 (i)-(ii) are standard in literature. Condition (iii) implies that the covariance of indicator functions between different observations of random variables X_i^c, Z_{mi}, Z'_{mi} and W_{mi} converges to zero as $n\to\infty$. The reason why we need this assumption is due to the structure of Z_{mi} (and Z'_{mi}), which contains individual's characteristics. Z_{mi} and Z_{mj} will be dependent with each other if they share same individual's characteristics. For example, if $Z_{mi}=(X_1+X_2+X_3)/3$ and $Z_{mj}=(X_3+X_4+X_5)/3$, then Z_{mi} and Z_{mj} are not independent because they share X_3. Similar argument applies to Z'_{mi} and Z'_{mj}. This condition ensures that standard law of large numbers works for these indicator functions.

Proposition 3.3. *Under Assumption 3.9,* $\widehat{f}_{Z_m|Z_m^*,X^c}\xrightarrow{p}f_{Z_m|Z_m^*,X^c}$ *and* $\widehat{f}_{Z_m^*|X^c}\xrightarrow{p}f_{Z_m^*|X^c}$ *for* $m=1,2$.

Next, we show that the second-stage pesudo-likelihood estimator for the copula parameter σ is consistent. Following Rio (2017), for real-valued random variables X and Y, we set

$$\tau(X,Y) = 2\sup_{(x,y)\in\mathbb{R}^2} |Pr(X > x, Y > y) - Pr(X > x)Pr(Y > y)|$$

as the strong mixing coefficient. The strong mixing coefficient between two σ-fields $\mathcal{A}$ and $\mathcal{B}$ is defined by

$$\tau(\mathcal{A},\mathcal{B}) = 2\sup\{|\operatorname{Cov}(\mathbb{1}_A, \mathbb{1}_B)| : (A,B) \in \mathcal{A}\times\mathcal{B}\}.$$

Let σ_0 denote the true value of σ. Also we use Σ and $\mathcal{F}_m$ denote the support for σ and $F_m \equiv F_{Z_m|X^c}$, $m = 1,2$, respectively. We impose the following assumptions,

ASSUMPTION 3.10. *(i) $\Omega \equiv \Sigma\times\mathcal{F}_1\times\mathcal{F}_2$ is a compact metric space with metric $\|\cdot\|$; (ii) For each $\sigma\in\Sigma$ and $F_m \in \mathcal{F}_m$, $m = 1,2$, $\log c_\sigma(F_1,F_2)$ is measurable; (iii) For each $Z_{mi} \in \mathcal{Z}_m$ and $X_i^c \in \mathcal{X}^c$, $\log c_\sigma(F_1,F_2)$ is continuously differentiable on Ω to order 3; (iv) $\lim_{n\to\infty}\tau(\mathcal{S}_i,\mathcal{S}_j) = 0$, where $\mathcal{S}_i$ denote the σ-field induced by (Z_i, Z_i', X_i^c, Y_i).*

Conditions (i)-(iii) are standard in the MLE literature. Condition (iv) is a stronger version of Assumption 3.9 (iii) and implies that the covariance between measurable functions of (Z_i, Z_i', X_i^c, Y_i) and (Z_j, Z_j', X_j^c, Y_j) will converge to 0 as $n\to\infty$ for $i \neq j$.

Proposition 3.4. *Under Assumptions 3.9-3.10, $\widehat{\sigma} \xrightarrow{p} \sigma_0$.*

Next, we establish the asymptotic normality of the semiparametric estimator $\widehat{\sigma}$ under suitable regularity conditions.

ASSUMPTION 3.11. *(i) $\sigma_0 \in \operatorname{int}(\Sigma)$; (ii) $b \equiv \mathbb{E}[\nabla_{\sigma,\sigma}\log C_\delta(F_1,F_2)] < 0$; (iii) For $m = 1,2$, $\|\widehat{F}_{Z_m,X^c} - F_{Z_m,X^c}\|_\infty = o_p(r_n)$, where $\|\cdot\|_\infty$ is the sup-norm and r_n is a nonstochastic positive real sequence such that $r_n < n^{-1/4}$ for each n; (iv) $\sum_{i=1}^n \sum_{j\neq i}^n \tau(\mathcal{S}_i,\mathcal{S}_j) < \infty$*

Conditions (i) and (ii) of Assumption 3.11 are standard in the literature. Condition (iii) requires that the uniform convergence rate of the empirical distribution function $\widehat{F}_{Z_m,X^c}$ to be faster than

$n^{-1/4}$, which is necessary for deriving the asymptotic distribution of semiparametric M-estimators, see, e.g., Newey (1994) and Newey and McFadden (1994). Condition (iv) is a technical condition that makes central limit theorem work for weakly dependent random variables.

Define

$$v^2 = \lim_{n\to\infty} \frac{1}{n} \mathrm{Var}\left\{ \sum_{i=1}^{n} [\nabla_\sigma \log C_\delta(F_1(Z_{1i}|X_i^c), F(Z_{2i}|X^c)) + K_1(Z_{1i}, X_i^c) + K_2(Z_{2i}, X_i^c) + K_3(X_i^c)] \right\},$$

where

$$K_1(Z_{1i}, X_i^c) = \int \mathbb{1}(Z_{1i} \le z_1, X_i^c \le x^c) \left[\nabla_{\sigma, F_{Z_1, X^c}} \log c_\sigma(F_1(z_1|x^c), F_2(z_2|x^c))\right] dF_{Z_1,Z_2,X^c}(z_1, z_2, x^c),$$

$$K_2(Z_{2i}, X_i^c) = \int \mathbb{1}(Z_{2i} \le z_2, X_i^c \le x^c) \left[\nabla_{\sigma, F_{Z_2, X^c}} \log c_\sigma(F_1(z_1|x^c), F_2(z_2|x^c))\right] dF_{Z_1,Z_2,X^c}(z_1, z_2, x^c)$$

and

$$K_3(X_i^c) = \int \mathbb{1}(X_i^c \le x^c) \left[\nabla_{\sigma, F_{X^c}} \log c_\sigma(F_1(z_1|x^c), F_2(z_2|x^c))\right] dF_{Z_1,Z_2,X^c}(z_1, z_2, x^c).$$

The asymptotic distribution of $\widehat{\sigma}$ can be summarized as follows.

Proposition 3.5. *Under Assumptions 3.9-3.11,* $\sqrt{n}(\widehat{\sigma} - \sigma_0) \xrightarrow{p} N(0, v^2/b^2)$.

Now, we show that the third-stage GMM estimator for the social parameter θ is consistent. Define $\eta = (f_{Z_1,Z_1',Y,X^c}, f_{Z_2,Z_2',Y,X^c}, f_{Z_1,Z_1',X^c}, f_{Z_2,Z_2',X^c}, f_{Z_1,X^c}, f_{Z_2,X^c}, f_{Z,X^c}, f_{X^c})^T$. Then, $g(Y_i, S_i, \theta)$ can be written as $g_i(\theta, \sigma, \eta)$. Let Θ and $\mathcal{H}$ denote the support for θ and η, respectively. Also use θ_0 to denote the true value of θ. We impose the following assumptions,

ASSUMPTION 3.12. *(i) Θ is compact; (ii) For each $\sigma \in \Sigma$ and $\eta \in \mathcal{H}$, $g_i(\theta, \sigma, \eta)$ is measurable; (iii) For each $Z_i \in \mathcal{Z}$, $g_i(\theta, \sigma, \eta)$ is continuously differentiable on $\Theta \times \Sigma \times \mathcal{H}$ to order 2; (iv) $\widehat{V} \xrightarrow{p} V$, a positive definite $(k+3) \times (k+3)$ matrix.*

Assumption 3.12 (i)-(iii) are similar to Assumption 3.10. Condition (iv) is standard in the GMM literature.

Theorem 3.2. *Under Assumptions 3.9, 3.10 and 3.12,* $\widehat{\theta} \xrightarrow{p} \theta_0$.

Next, we establish the asymptotic normality of the semiparametric GMM estimator $\widehat{\theta}$ under suitable regularity conditions.

ASSUMPTION 3.13. *(i)* $\theta_0 \in \mathrm{int}(\Theta)$*; (ii)* $G \equiv \{\mathbb{E}[\nabla_\theta g_i(\theta_0, \sigma_0, \eta_0)]\}^T V \mathbb{E}[\nabla_\theta g_i(\theta_0, \sigma_0, \eta_0)]$ *is positive definite; (iii)* $\|\widehat{\eta} - \eta_0\|_\infty = o_p(r_n)$*, where* r_n *is a nonstochastic positive real sequence such that* $r_n < n^{-1/4}$ *for each n.*

Assumption 3.13 is similar to Assumption 3.11 (i)-(iii) and is standard in the literature. First, we introduce some notations. Define $R = \mathbb{E}[\nabla_\theta g_i(\theta_0, \sigma_0, \eta_0)]$, $\zeta_i = (Z_i, Z_i', Y_i, X_i^c)$, $\zeta_{mi} = (Z_{mi}, Z_{mi}', Y_i, X_i^c)$, $\widetilde{\zeta}_i = (Z_i, Z_i', X_i^c)$, $\widetilde{\zeta}_{mi} = (Z_{mi}, Z_{mi}', X_i^c)$,

$$H_{1m}(\zeta_{mi}) = \int \mathbb{1}(\zeta_{mi} = \zeta_m) \nabla_{f_{Z_m, Z_m', Y, X^c}} g_i(\theta_0, \sigma_0, \eta_0) dF_{Z,Z',Y,X^c}(\zeta),$$

$$H_{2m}(\widetilde{\zeta}_i) = \int \mathbb{1}(\widetilde{\zeta}_i = \widetilde{\zeta}_m) \nabla_{f_{Z_m, Z_m', X^c}} g_i(\theta_0, \sigma_0, \eta_0) dF_{Z,Z',Y,X^c}(\zeta),$$

$$H_{3m}(Z_{mi}, X_i^c) = \int \mathbb{1}\left(Z_{mi} = z_m, X_i^c = x^c\right) \nabla_{f_{Z_m, X^c}} g_i(\theta_0, \sigma_0, \eta_0) dF_{Z,Z',Y,X^c}(\zeta),$$

$$H_{4m}(Z_i, X_i^c) = \int \mathbb{1}\left(Z_i = z, X_i^c = x^c\right) \nabla_{f_{Z, X^c}} g_i(\theta_0, \sigma_0, \eta_0) dF_{Z,Z',Y,X^c}(\zeta),$$

$$H_{5m}(X_i^c) = \int \mathbb{1}\left(X_i^c = x^c\right) \nabla_{f_{X^c}} g_i(\theta_0, \sigma_0, \eta_0) dF_{Z,Z',Y,X^c}(\zeta),$$

and

$$\Lambda = \lim_{n \to \infty} \frac{1}{n} \mathrm{Var} \left\{ \sum_{i=1}^{n} [g_i(\theta_0, \sigma_0, \eta_0) + \nu_i + \xi_i] \right\},$$

where

$$\nu_i = \mathbb{E}\left[\nabla_\sigma g_i(\theta_0, \sigma_0, \eta_0)\right] \frac{\nabla_\sigma \log C_\delta(F_1(Z_{1i}|X_i^c), F_2(Z_{2i}|X^c)) + K_1(Z_{1i}, X_i^c) + K_2(Z_{2i}, X_i^c) + K_3(X_i^c)}{-\mathbb{E}[\nabla_{\sigma,\sigma} \log C_\delta(F_1, F_2)]}$$

and

$$\xi_i = \sum_{m=1}^{2} \left[H_{1m}(\zeta_{mi}) + H_{2m}(\widetilde{\zeta}_i) + H_{3m}(Z_{mi}, X_i^c) + H_{4m}(Z_i, X_i^c) + H_{5m}(X_i^c) \right].$$

The asymptotic distribution of $\widehat{\sigma}$ can be summarized as follows.

Theorem 3.3. *Under Assumptions 3.9-3.13,* $\sqrt{n}(\widehat{\theta} - \theta_0) \xrightarrow{d} N(0, G^{-1}R^T V \Lambda V R G^{-1})$.

3.4 Empirical Study

In applied work, firm linkages are often constructed based on if two firms are in the same industry. However, it is unlikely to have every firm in the same industry connected. Besides, for different kinds of industry classification code(e.g. SIC industry code, GIC code, Standard Poor industry classification, etc.), classifications on firms are different, as a result the associated firm networks are different. Other applied work on firm networks use data that to some extent describe the relationship between two firms (e.g. customer-supplier data, connection data of members in the boards). However, for a static network study, such data may be just a instantaneous observation on the true network of interest. There can be firms that are linked but unobserved and also firms that are only interacted for a very short period but still under observed. Therefore, in the study of firm networks, the problem of mismeasured links is quite commonly seen, hence requires a lot attention to deal with.

In this section, we carry out an empirical analysis to study the peer effects among firms on their corporate financial policies. We construct both peer-effects network and contextual-effects network based on a sample of 419 firms. The peer-effects network is constructed by defining the linkage between two firms based on if they form a customer and supplier link. It has been seen in the literature that corporate financial policies made by customers do affect corporate financial policies made by suppliers, and vice versa. The contextual-effects network is constructed by defining the linkages between two firms based on if they are in the same industry. In general, firms that in the same industry have similar characteristics and these similarities have impact on firms' corporate financial policies. Empirical findings indicate that the peer effect and contextual effect among firms are statistically significant. But if we ignore the existence of mismeasured network links, peer effect becomes insignificant.

3.4.1 Data and Summary Statistics

Our primary data come from two sources: the Customer-supplier data for the period 2001 and 2002, and the merged Center for Research in Security Prices (CRSP)-Compustat database for the period 1996 to 2001. After merging two data sets and getting rid of isolated firms in customer-supplier data, we get the full sample with 419 observations.

The dependent variable is the leverage ratio chosen by each firms. Leverage ratio is defined as the ratio between a firm's debt and equity, which is regarded as a choice of firm's financial policy. The independent variable is return shock, introduced by Leary and Roberts (2014), that is constructed to be strictly exogenous. We follow Leary and Roberts (2014) regression procedure to predict this idiosyncratic shock in 2001 by using data from period 1996 to 2000. Details on obtaining return shock is discussed in Section 7.1.2.

We allow for four variables to control for the firms' capital structures and profitability. These variables are firms' market to book ratio, logarithm of sales turnover rates, the ratio between earnings per shares from operations and assets, and the ratio between net property, plant, and equipment and assets. The market to book ratio and the ratio between net property, plant, and equipment and assets are two variables that describe a firm's capital structure. The logarithm of sales turnover rates and the earnings per shares from operations control for a firm's ability in making profit. In our identification method, an instrument variable as indicator is required. We use firm size, defined as logarithm of assets, as the indicator variable (Z'').

Table 12 provides the summary statistics on control variables, the dependent variable, the independent variable and the instrument variable.

We use the Customer-supplier data to construct firm's peer-effect networks (A and A'), by

Table 12: Summary Statistics of Variables

Variable	Mean	Std. Dev.	Min.	Max.	Sample Size
Control variables					
Market to book ratio	2.5208	7.9466	-33.2166	121.874	419
Log(sales/turnover)	1.9564	1.1426	-1.4815	4.6296	419
EPS from operations/assets	-0.0758	0.7337	-8.9888	4.7416	419
Net PPE/assets	0.2472	0.1942	0	0.9743	419
Dependent variable					
Leverage	0.2947	3.765	-46.811	46.2578	419
Independent variable					
Idiosyncratic return	0.0315	0.2724	-0.6007	3.1792	419
Instrument variable					
Log(Assets)	6.1885	2.5634	0.6632	13.3403	419

defining the firms as the nodes and the customer-supplier pairs as the links. The customer firms' choices on financial policies can affect the supplier firms' choices of financial policies, and vice versa. So the adjacency matrices for firms' peer-effect networks are undirected and symmetric. Specifically, the Customer-supplier data in 2001 is used as the first measurement for peer-effect network (A); while the Customer-supplier data in 2002 is used as the second measurement (A').

We get the firms' contextual-effect networks (C and C') based on the industry classification codes. The nodes are defined as the same set of firms in the Customer-supplier data and the links are defined as the firm-pairs that are from the same industry. Firms' choices on financial policies can be affected by the characteristics of other firms that are from the same industry. Specifically, we use the 3-digit SIC code as the first measurement for contextual-effect network(C) and the GIC industry code as the second measurement (C').

Figure 6 presents the matrix plots of the same 50 firms from our sample. We found for both the contextual-effects network and the peer-effects network, the repeated measurements indeed have different link structures. Besides, a larger fraction of links are the same in the repeated measurements. Figure 7 provides the histogram on degrees of all measurements of the peer-effects network and the contextual-effects network. We require no isolated firms in peer-effect networks but allow for isolated firms in contextual-effect networks. From Figure 7, we observe that all the four networks, especially the peer-effect networks, are sparse, as required by our model.

In our model, we require X_i to be an exogenous peer firm characteristic. It is not trivial to find such exogenous characteristic even by controlling for firm i's own characteristics. Following the identification strategy in Leary and Roberts (2014), we begin with a known capital structure determinant, stock returns, for the time period Jan.1991 to Dec. 2000. We use a traditional asset pricing model to extract the idiosyncratic variation, along with the variation among peers in stock returns. We obtain the residual from this model and define this residual as the return shock and use it to obtain the exogenous variation in peer firms' characteristics in year 2001. Such exogenous variation in year 2001 is used as X_i in our model.

The intuition behinds the return shock follows the event study approach. We need to identify certain events that are relevant for peer firms but are random conditional on observables- firm i's capital structure. Such events or shocks can be accidental CEO deaths, accounting scandals, and so on. But we cannot directly apply these events because such events are rare to see and lack of external validity. Besides, it is also unclear whether these events are exogenous because of spillover effects. Therefore, we need to address the shortcomings in event study approach. The details on obtaining return shock are discussed as below.

First, we run a regression for each firm's stock return on a rolling basis using historical monthly data for period from January 1996 to Decemeber 2000:

$$r_{ijt} = \alpha_{ijt} + \beta_{ijt}^{M}(rm_t - rf_t) + \beta_{ijt}^{IND}(\bar{r}_{-ijt} - rf_t) + \eta_{ijt}, \tag{3.17}$$

where r_{ijt} refers to the stock return for firm i in industry j over month t, $(rm_t - rf_t)$ is the monthly excess market return, and $(\bar{r}_{-ijt} - rf_t)$ is the excess return on an equal-weighted industry portfolio excluding firm i's return. We use the first measurement for contextual-network as such industry classification. The idiosyncratic return is given by:

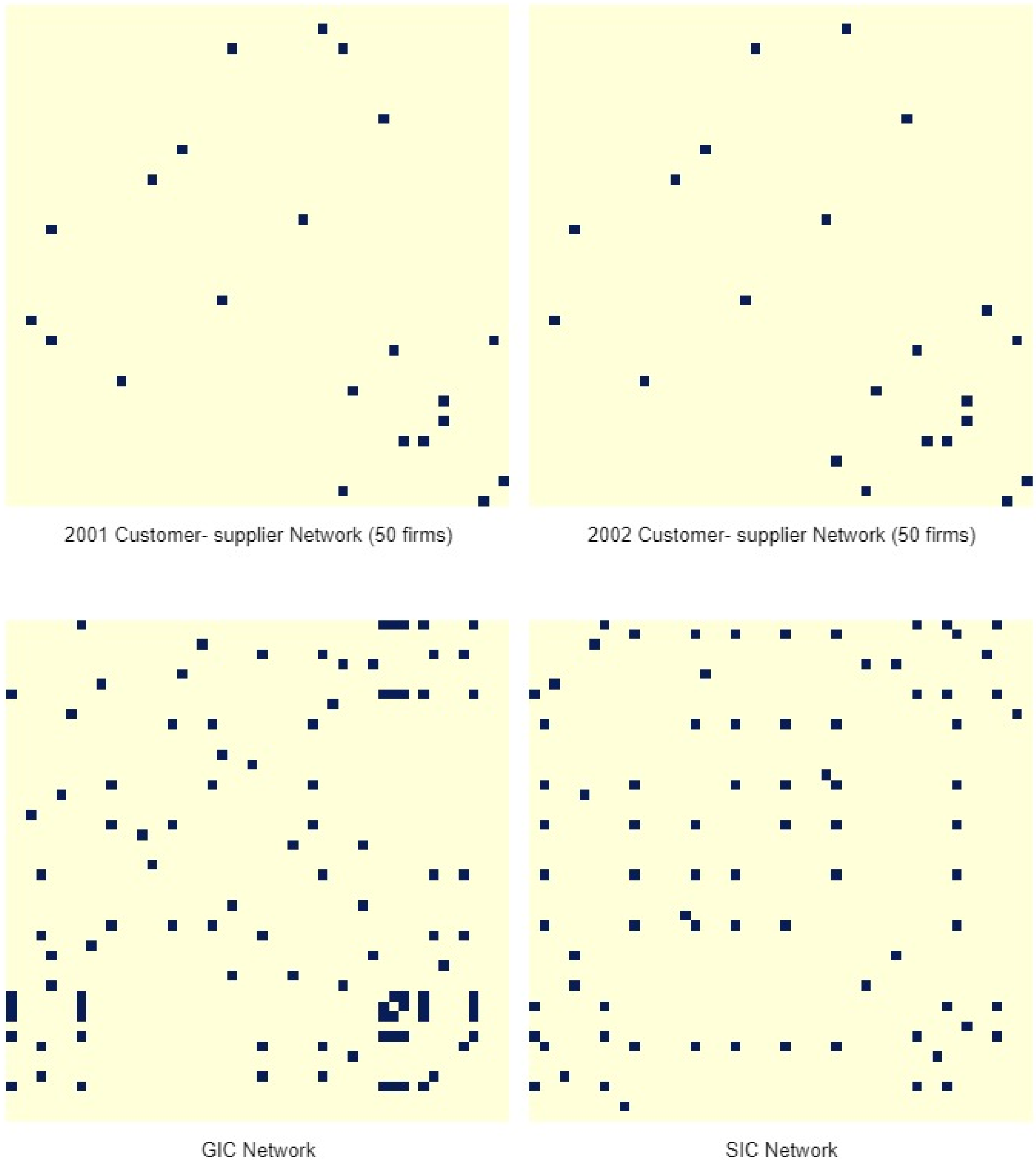


Figure 6: Networks for Fifty Firms

Figure 6 presents the matrix plots of the same 50 firms from our sample. We found for both the contextual-effects networks and the peer- effects networks, the repeated measurements indeed have different link structures.

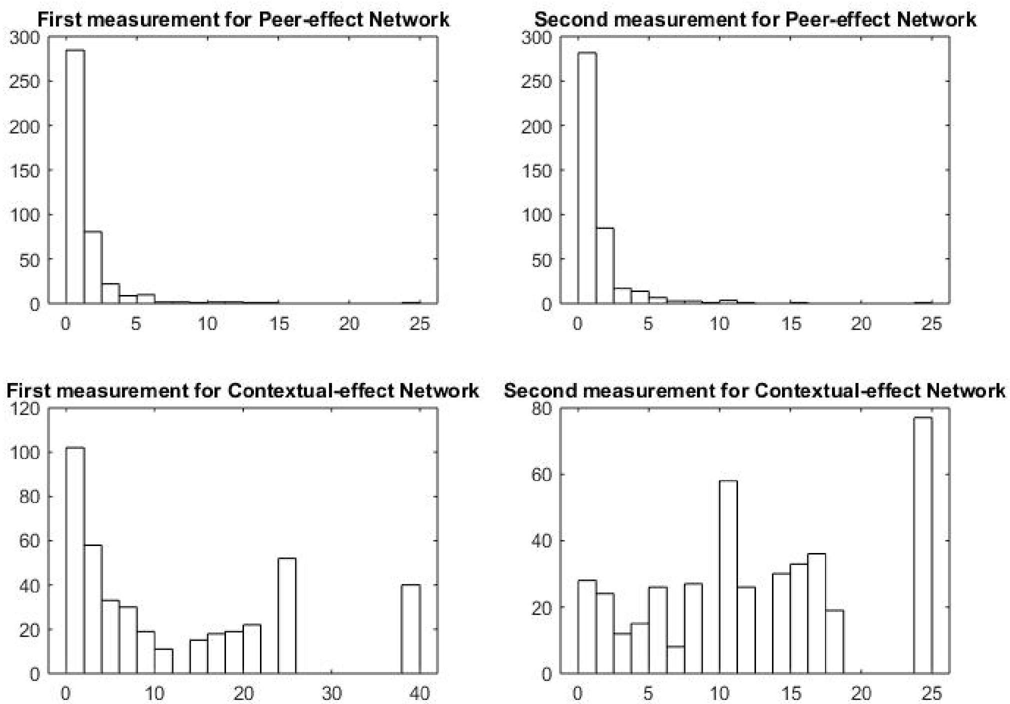


Figure 7: Histograms of Degree in Each Network

Figure 7 presents the histogram figures of number of links in each network. The horizontal axis is the degree level, which measures the number of direct links of a firm. The vertical axis is the number of firms.

$$Idiosyncratic\ Return_{ijt} \equiv \hat{\eta}_{ijt} = r_{ijt} - \hat{r}_{ijt} = \hat{\alpha}_{ijt} + \hat{\beta}^{M}_{ijt}(rm_t - rf_t) + \hat{\beta}^{IND}_{ijt}(\bar{r}_{-ijt} - rf_t) \quad (3.18)$$

We require at least 20 months and up to 60 months of historical data for each firm in the estimation. For example, to obtain the idiosyncratic returns for Google Inc. between January 2001 and December 2001, we first estimate equation 3.17 using data between January 1996 and December 2000. Once we get the estimators $(\hat{\alpha}_{ijt}, \hat{\beta}^{M}_{ijt}, \hat{\beta}^{IND}_{ijt})$, we use these parameters to predict the expected return in 2001 using historical data on excess market returns and excess industry returns for Google Inc. in 2001. Then as shown in equation 3.18, we subtract the expected return from the Google Inc. return in 2001 to get the idiosyncratic return for Google Inc. in year 2001.

Table 13 presents summary statistics for the return shock regression results. We get estimation results for each firm in the sample, so the sample size for each result is also 419. The mean of number of observations per regression is 37 and the average adjusted R^2 is 13%. The average coefficients on market excess return and excess industry return are both positive. The average expected return is 0.0349 and the average idiosyncratic return is 0.0315. The 419 predicted results of idiosyncratic return is used as X_i, for each i in our model.

Table 13: Summary Statistics of Return Shock Regression

Variable	**Mean**	**Std. Dev.**	**Min.**	**Max.**	**N**
Summary statistics of variables					
Stock returns	0.0665	0.274	-0.4545	3.2727	419
Return excess market return	-0.0509	0.0579	-0.1424	0.0579	419
Return excess avg. industry return	0.0195	0.1517	-0.3131	0.6429	419
Summary statistics of regression results					
α_i	0.0414	0.0365	-0.083	0.1582	419
β_i^M	0.3676	0.6930	-1.8269	2.6499	419
β_i^{IND}	0.4521	0.5212	-0.9703	3.9235	419
Obs. per regression	37.0406	11.5934	20	60	419
Adjusted R^2	0.1347	0.1467	-0.1023	0.7023	419
Summary statistics of predicted results					
Expected return	0.0349	0.1153	-0.2659	0.5279	419
Idiosyncratic return	0.0315	0.2724	-0.6007	3.1792	419

3.4.2 The Empirical Model

Our empirical model is a generalization of a linear-in-mean model, where we distinguish the peer-effects network with the contextual-effects network,

$$Y_i = \alpha + \gamma' X_i^c + \beta \sum_j A^*_{ij} Y_j + \delta \sum_j C^*_{ij} X_j + \varepsilon_i,$$

where the index i and ij correspond to firm i and links between firms i and j, respectively. The outcome variable, Y_i, is a measure of corporate financial policy, such as leverage. The variable $\sum_j A^*_{ij} Y_j$ denotes the average of peer firm leverage ratios in the peer-effects networks. Since all the diagonal elements in the adjacency matrix A^* are zeros, such average leverages exclude self-leverage. The K-dimensional vectors X_i^c are control variables including measures for firm capital structure and measures for firm profitability. The key independent variable, X_i, is the return shocks that are exogenous.

We refer to the parameter vector of interest, $(\alpha, \gamma, \beta, \delta)$, as structural parameters, to distinguish them from reduced-form parameters that appear in the context of instrumental variables. Social effects are captured by β and δ, which measure the peer effect and the contextual effect, respectively, on firms' leverage ratios. Under our identification strategy, with the help of instrument variables, we obtain the reduce form equation,

$$Y_i = \tilde{\alpha} + \tilde{\gamma}' X_i^c + \tilde{\beta} \sum_j A^*_{ik} C^*_{kj} X_j + \tilde{\delta} \sum_j C^*_{ij} X_j + \varepsilon_i,$$

where the notation tilde refers to reduced form parameters that are functions of the structural parameters. Specifically,

$$\tilde{\alpha} = \frac{\alpha}{1-\beta};\ \tilde{\gamma} = \gamma;\ \tilde{\delta} = \delta;\ \tilde{\beta} = \delta\beta.$$

Our empirical estimation follows the estimation strategy introduced in Section 5. In the first step, we obtain the empirical joint distribution of the first measurement and the second measurement of the mean characteristics(Z_{1i} and Z'_{1i}), and the empirical joint distribution of the first and the second measurement of the instrumental mean characteristics(Z_{2i} and Z'_{2i}). Also, we obtain the above empirical joint distributions conditional on the instrument indicator variable, Z'', which is an indicator of firm size. We use the logarithm of firm assets and set it as a binary variable. Z'' is one if firm asset is greater than 75 (in millions of dollars), and zero otherwise. Then we use an eigenvector decomposition to get the predicted conditional distribution, $f_{Z''=z}(Z_1^*|Z_1)$ and $f_{Z''=z}(Z_2^*|Z_2)$, where $z \in \{0,1\}$. Moreover, we are able to obtain the predicted distributions of the latent mean characteristics ($D_{Z_1^*}$) and the latent instrumental mean characteristics ($D_{Z_2^*}$).

The top two sub-figures of Figure 8 present the empirical cdfs of Z_1 and Z_2 as observed from data. The blue lines present the empirical cdfs for the two measurements of the mean characteristics. The orange lines present the empirical cdfs for the two measurements of the instrumental mean characteristics. From the empirical densities of observed data, we find that Z_1 and Z_2 follows similar patterns of cumulative densities. The bottom two sub-figures of Figure 8 present the empirical cdfs of discretized Z_1 and Z_2 based on observed data. To get the best fit discretization of Z_1 and Z_2, we check for the minimum condition numbers of joint distribution of Z_1 and Z_2. We find that the discretized empirical cdfs follow similar patterns as the observed empirical cdfs.

Figure 9 presents the histograms of the distirbution of the predicted latent (Z_1^*, Z_2^*) and the observed (Z_1, Z_2). We find that the probability density of the lower mean characteristics is much larger than that of the observed mean characteristics. The mean of the latent mean characteristics is less than the observed mean. This indicates in the latent contextual-effects network, there are more links between firms with lower return shocks rather than those between firms with higher return shocks. The mismeasured links mostly lie in the fact that we observe absent links to higher return shock firms, while do not observe existing links to lower return shock firms.

In the second step, we use the empirical marginal distribution of Z_1 and Z_2 to estimate the parameter of Gumbel copula function (a copula from Archimedean Family). We transform the data

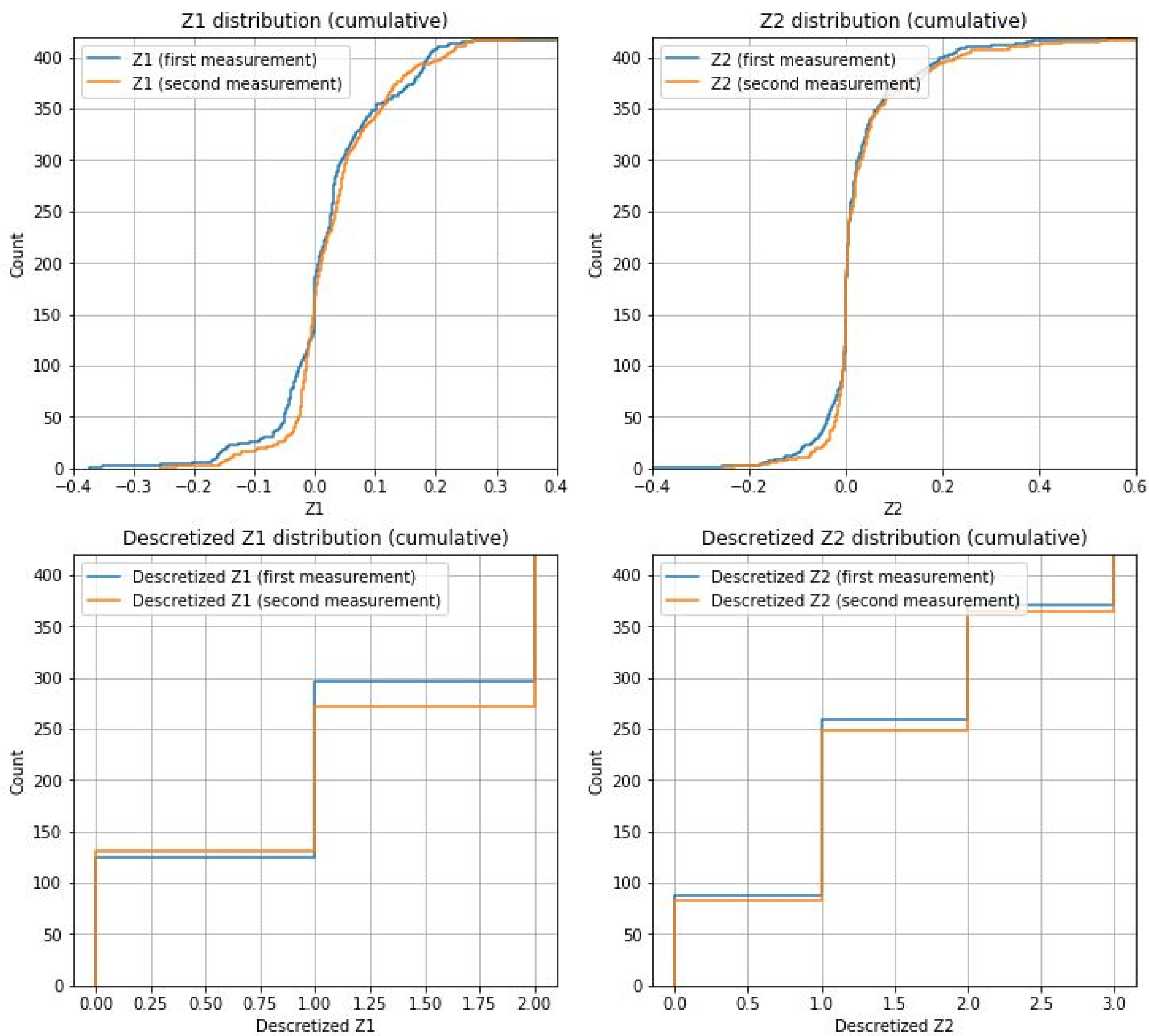


Figure 8: Empirical CDF of Z and Discretized Z

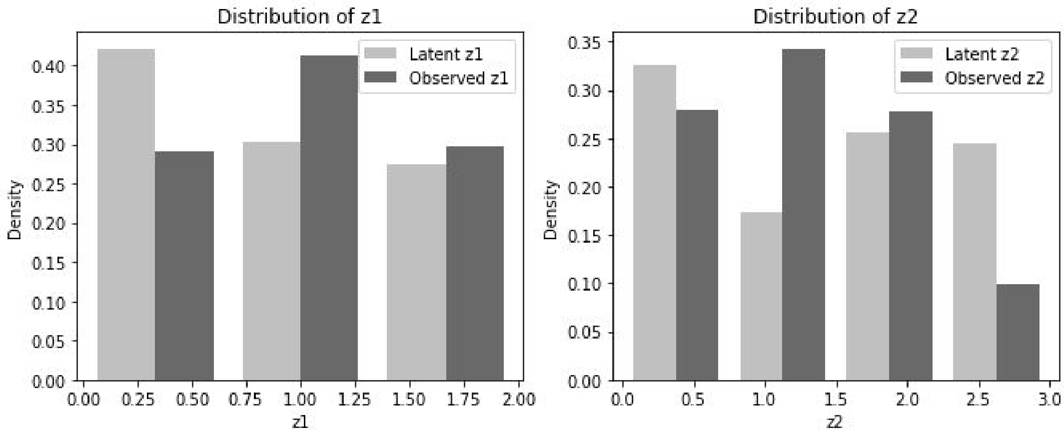


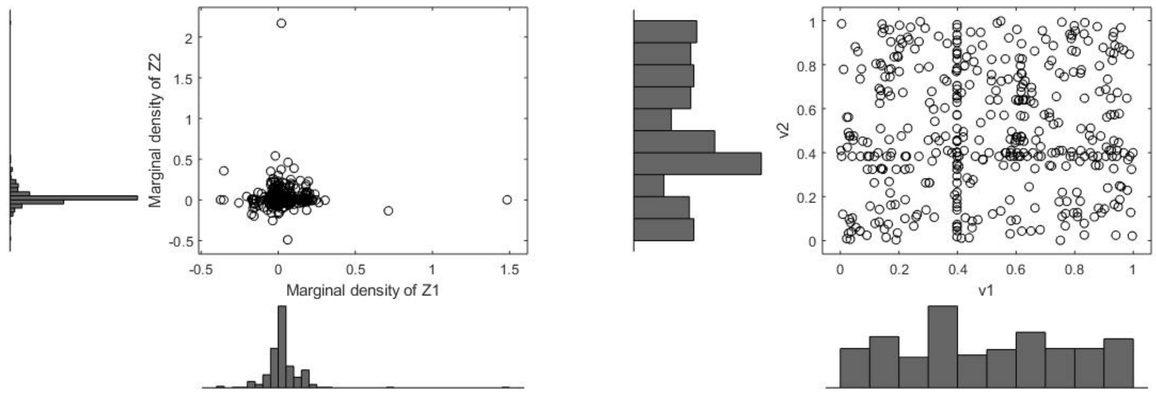
Figure 9: Distributions of Latent Z and Observed Z

to the copula scale (unit square) using a kernel estimator of the cumulative distribution function. Figure 10a presents the marginal density of Z_1 and Z_2 respectively in histograms and joint density of Z_1 and Z_2 in circle plots. Figure 10b presents the densities after transformation using a kernel estimator of cdf. Figure 11 presents the pdf and cdf of the Gumbel Copula of Z_1 and Z_2 using the estimated copula parameter. We use this estimated parameter, along with Gumbel copula function to obtain the joint distribution of Z_1^* and Z_2^*.

In the third step, we use the predicted distributions to recover the latent mean characteristics and the latent instrumental mean characteristics for each firm. Then we use a GMM method to recover the reduced-form parameters and their associated statistics. We transform the reduced-form parameters into structural parameters and obtain the associated statistics by using a bootstrap method.

3.4.3 Empirical Results

Table 14 presents the structural estimates of the GMM estimation. The standard errors in parentheses are obtained using bootstrap method. We find that both the latent peer effect and the latent contextual effect are significant and positive. If the average leverage ratio of a firm's customer/supplier is increased by 1 unit, the firm will also increase its leverage by 0.8727. The positive peer effects



(a) Marginal Density of Z1 and Z2 (b) Transformed Density of Z1 and Z2

Figure 10: Copula Transformations

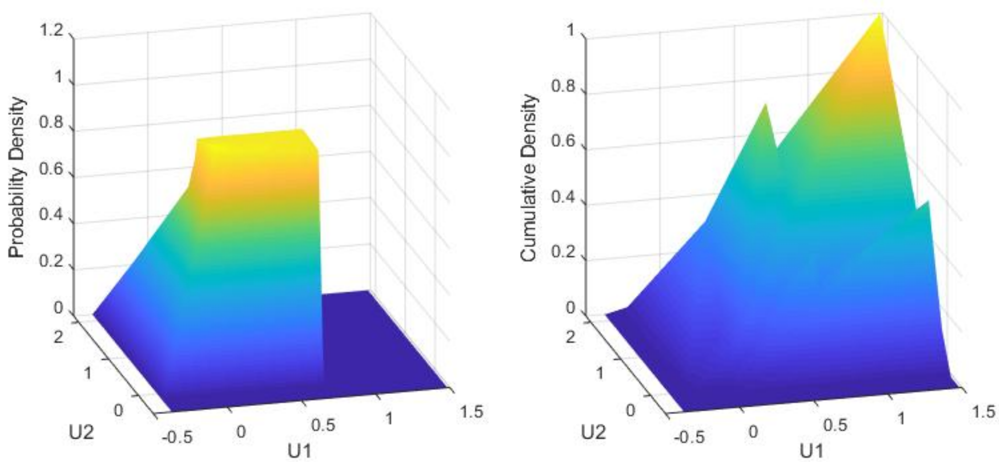


Figure 11: Estimated PDF and CDF of a Gumbel Copula of Z1 and Z2

on leverage ratios are found in the literature. Firms often have very close connections with their key suppliers. Indicated by the bargaining power theory, when firms raise their leverages, their suppliers or customers will also raise their own leverages in response to increase their bargaining power (Hennessy and Livdan (2009), Chu (2012)). Our estimation result also indicates a positive contextual effect. In the same industry, increased return shock to peer firms implies an industry prosperity. In turn, firms choose to raise leverages in order to take chance in such good industry environment. We also find positive effect of earnings per share from operations and assets on leverage ratio. Earnings per share is a measure on firm's profitability with higher EPS indicating higher ability in running business. This implies a firm with better profitability will choose higher leverage. Market-to-book ratio, a measure on firm's capital structure, was found to have a negative effect on leverage ratio.

We also conduct a specification by estimating the structural parameters in the case where we observe networks without any mismeasured links. Table 15 presents the GMM results. The peer effects and the contextual effects are both insignificant, and also with lower magnitude compared with the case where we consider measurement errors.

Empirical findings indicate that if we ignore the mismeasured links in firm networks, the estimate results are underestimated and most importantly, insignificant. After we address the concern of mismeasured links, peer effects are found to be significant, with a larger magnitude. Such change in estimate results implies importance in being cautious about mismeasured links. If the data we use are self-reported, or only for a short time span, or the method to construct a network has some embedded problems in defining network links, such as the industry code, then it is important to consider the chance having mismeasured links in the observed network. Since failing to deal with these mismeasured links, by using the observed network as the true network of interest, can result in biased estimates of peer effects parameters as indicated by our empirical analysis, or can result in other undesired consequences as introduced in the previous sections.

Table 14: Structural Parameter Estimation

	Estimated Results
$\sum_j C^*_{ij} X_j$	0.3479*
	(0.2057)
$\sum_j A^*_{ij} Y_j$	0.8727**
	(0.4011)
Market-to-book ratio	-0.1485*
	(0.0821)
Log(sales/turnover)	0.0572
	(0.0643)
EPS from operations/assets	0.0659*
	(0.0398)
Net PPE/assets	0.0727
	(0.0912)
Constant	-0.0936
	(0.1280)
Sample Size	419

standard errors in parentheses

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 15: Parameter Estimation: Mismeasured Links are Ignored

	Estimated Results
$\sum_j C^*_{ij} X_j$	0.2658
	(0.2026)
$\sum_j A^*_{ij} Y_j$	0.2415
	(0.1895)
Market-to-book ratio	-0.1428*
	(0.0770)
Log(sales/turnover)	0.1379
	(0.1533)
EPS from operations/assets	0.0108
	(0.0569)
Net PPE/assets	-0.0013
	(0.0319)
Constant	0.0272
	(0.0677)
Sample Size	419

standard errors in parentheses

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

3.5 Conclusion

Peer effects are widely studied for various aspects of network outcomes. Previous literature on identification of peer effect parameters mainly require the knowledge of adjacency matrices. In this paper, we show the sufficient conditions to identify and estimate the peer effects in networks where links are latent because of measurement errors. We can draw two main conclusions by adopting our estimation method to the firm networks. First, we find that both the peer effects and the contextual effects are significantly positive among firms on the decisions of leverage ratios. Second, we show that the peer effects are not significant if the measurement errors are ignored. There could be two feasible and enlightening tasks to be considered as part of the future works. One can think of using weighted adjacency matrix instead of unweighted adjacency matrix to characterize firm networks under the concepts of sale amounts between customers and suppliers. By weighting the adjacency matrix with different importance in linkages, one can tell whether a pair of customer and supplier are strongly connected because they have large amount of sales happened. Also, one can think of discrete choice outcome variable instead of continuous outcome variable. For example, if one are interested in the peer effects on firm's R&D choice, the binary outcome variable requires a new way to identify the peer effects.

4 FIRST PRICE AUCTION WITH AFFILIATION: A NETWORK PERSPECTIVE

Auction with affiliated private values (APV) are widely studied (Milgrom and Weber (1982), Li, Perrigne and Vuong (2002), Hubbard, Li and Paarsch (2012), Li and Zhang (2015)). Previous literature on first-price sealed-bid auction with affiliation mainly require symmetric dependent structure among all bidders. However, in some applied work, it is unrealistic that every possible combination pairs of bidders have the same dependent structure due to information asymmetry. This paper conducts a reduced-form analysis of first-price sealed-bid auction with affiliation under a network perspective. Specifically, we provide some empirical evidence on the dependency of private values among bidders based on their linked status and proposes a structural model for estimation.

The objective of this paper is to empirically study the first-price sealed-bid auction with affiliation, given the network of potential bidders. In the existing literature, researchers generally adopt symmetric dependency among bidders' private values through affiliation (Li and Zhang (2015)). Nevertheless, how the dependency determined by link status between bidders remains unclear.

To investigate the the first-price sealed-bid auction with affiliation under a network perspective, we take several steps. In the first step, we propose a simple model that characterize our basic settings and assumptions. For the isolated bidders, their private values are independent of others'. For the linked pairs of bidders, their private values are dependent. We further assume that the dependence structure of private valuations between any pair of linked bidders is the same, regardless of the auction and bidder characteristics. Basically, this model introduces asymmetry into the affiliation relationship between bidders, where such asymmetry is fully captured by their network structure. The reduced-form analysis are all based on the idea of dependency between linked pairs.

Next, we construct contract-specific contractor networks using data from California Department of Transportation (Caltrans). The contract-specific contractor network is constructed by the chance of having at least one same subcontractors in the underlying contract. Intuitively, we expect that a subcontractor can effectively serve as a go-between that makes information (e.g. costs, qualities, working days, etc.) sharing available for two contractors. According to the regulation of Caltrans bidding process, the contractor must submit its bid along with a list of subcontractors. For a same project (contract), if two contractors listed at least one same subcontractor, their expectations on unit prices of project items should be close. As a consequence, the total bid amount submitted by two contractors should be close too.

Then based on the 712 contract-specific contractor networks from our sample, we investigate the densities of those networks. Density is a network descriptive measure that equals the number of links divided by the maximum number of potential links in a network. Our affiliation model requires that if two contractors are linked, then their private values are dependent; If two contractors are unlinked, then their private values are independent. Hence, density is very straightforward in telling us to what extent private values are dependent among contractors in a contractor network. The symmetric dependency structure among all pairs of bidders required by Li and Zhang (2015) only coincides with the case where density equals 1. In other words, it requires the network is a complete network with all nodes to be linked with each other. We find that 66.22% of our contractor networks are not complete networks, and even 16.49% of them are empty networks, indicating a quite different starting point compared with Li and Zhang (2015).

Moreover , we provide key evidence on how private values are dependent based on network structure. We first introduce the distance between two contractors and we find most pairs of contractors are either unlinked or within two-step reach. Then we construct a variable called Bid Amt, a ratio between total bid amount and engineer estimates. We investigate how the difference in Bid Amt between two contractors are determined by the distance between the two contractors. Besides that, we define a variable called Bid Gap as the difference on total bid amount between the winner and the biggest loser, divided by engineer estimates. We run a linear regression of Bid Gap on contractor-network density, controlling for the number of contractors in each contract. Intuitively, a network with higher density should have smaller bid gap due to higher chances of information exchange.

Lastly, we examine how the position of a contractor in its contractor network may affect its bidding behaviour and its probability of winning. We first run a linear regression of Bid Amt on four centrality measures (from social network analysis: degree, eigenvector, betweenness, closeness). Then we conduct a Probit model of probability of winning on centrality measures.

Empirical findings indicate that the effect of directly linked status on difference in Bid Amt is significantly negative with magnitude 0.0681 at 99% confidence level. This implies that if two contractors are linked in its contractor network, then they are likely to submit a similar bid price compared with unlinked or indirectly linked contractors. The result from specification on second-degree linked tells us that the indirect links have zero effect on Bid Amt. Intuitively, second-degree link involves two different subcontractors that brings possibly quite different information on bids. By studying Bid Gap and network densities, we find that the effect of density on Bid Gap is also negative at 99% significance level. In the case where network density changes from 0 to 1, the bid gap decreases by 0.209 significantly and the unit bid gap (controlling for the number of contractors) decreases by 0.0339 significantly. Additional empirical results related to network centralities implies the existence of network effects among contractors on their bidding amounts and probabilities of winning.

This paper contributes to the existing literature of auction with affiliated private values by providing a network perspective on the way of affiliation. In the existing literature, the dependency of private values among bidders are assumed to be symmetric (Li, Perrigne and Vuong (2002), Hubbard, Li and Paarsch (2012)). In Li and Zhang (2015), their model requires all the possible combination of two bidders to have exactly the same dependent structure. In our model, the dependency comes from the fact that two contractors are linked in its contractor network. If two contractors are unlinked, then their private values are considered independent.

The rest of the paper is organized as follows. Section 4.1 introduces the data source and summary statistics. Section 4.2 formally proposes our model. Section 4.3 introduces the construction of contractor networks and network properties. Section 4.4 studies the dependency of bid amounts based on link status. Section 4.5 investigates how network structure may affect bid amount and probability of winning. Section 4.6 discusses conclusions and potential future works.

4.1 Data and Summary Statistics

The primary data we use are collected from the detailed bid summary files provided by the California Department of Transportation (Caltrans). Each original PDF file summarizes a sealed-bid first-price auction with information on bidders (contractors), bid amount, contract characteristics and subcontractors.

At the beginning of bidding process, Caltrans advertises a project with information on bid open date, a list of items that required to complete the project, along with engineer estimates on the quantity and cost of each contract item. The potential bidders need to submit its sealed-bid on a unit price of each unit contract item, and the total bid amount will be the sum across all contract items by quantity. In the sealed-bid, the contractors must list each subcontractor whose work accounts for at least 0.5% or $10,000, whichever is greater, of the contract value. Each subcontractor must be prequalified to do the listed work. In the end, the contractor with the lowest total bid amount is awarded the project.

Our data collects bid summary in year 2001 and 2002, with a sample size of 14,983. For each unit of observation, we observe contract ID, contract open date, total amount of engineer estimates on the project, a contractor that submitted a bid amount upon some contract, along with one of the bidder's subcontractor. Notice that there can be multiple contracts per day, multiple bidders per contract and multiple subcontractors per bidder per contract. For convenience, we assign contracts, contractors, and subcontractors with unique numeric IDs based on their names. During this period, in summary, there are 712 construction contracts out for bid, 604 unique bidders ever joined at least one bid and 1723 unique subcontractors that ever worked with at least one of those bidders. Detailed summary statistics for the number of contracts, bidders and subcontractors can be found in table 16. The average number of contractors per contract is 5 and the average number of subcontractors that ever worked with one single contractor is 15.

Table 16: Summary Statistics: Bidders and Subcontractors

Variable	Mean	Std. Dev.	Min.	Max.	N
Number of Bidders					604
Number of Subcontractors					1723
Number of Contracts					712
Number of Contracts per Day	6.35	5.58	1	45	
Number of Bidders per Contract	5.05	3.05	1	22	
Number of Subcontractors per Bidder	15.00	19.73	1	278	

Summary statistics on contract characteristics are presented in Table 17. Before the Department of Transportation in California advertised a project (a contract), an estimate amount on the quantity and unit cost of the items that are required to complete the whole construction procedure was given by some technical engineer. We denote this estimated amount for bidders' bidding reference as Engineer Estimates. The total amount of engineer estimates varies a lot based on different types of project. The smallest contract only has an estimates of $78,000, while the largest contract has an estimates of $105,610,000. Amount Under Estimates is a variable that shows how much lower a contractor bids compared with the engineer estimates. Contractors are aware of the Number of Items in each contract before bidding. The average number of items per contract is about 51. The variable Working Days is the total working days spent by the winning contractor. Bid amount is the total bid amount on each project. For each bidder that submitted a bid for one project (contract), we observe a group of subcontractors that are listed to work on different aspects of contract items. For example, on Jan.4. 2001, there were three bidders bidding for a contract 01-411304: Mendocino Construction, Parnum Paving and Rolling rock Construction. Mendocino had three subcontractors: Anrak corporation, Central striping, and G and L Traffic. Parnum Paving had one subcontractor: Central Striping. Rolling rock had two subcontractors: Central Striping and G and L Traffic. In the end, Parnum Paving won the contract with a bidding amount of $1,379,983 and it finished the construction using 70 days.

Table 18 presents regression results of Bid Amount on three contract characteristics: Engineer

Table 17: Summary Statistics: Contract Charateristics

Variable	Mean	Std. Dev.	Min.	Max.	Sample Size
Engineer Estimates	5477825	13759460	78000	105610000	14983
Amount Under Est	670791	2510059	-7457202	21579142	14983
Number of Items	51	54	1	326	14983
Working Days	223	286	10	1350	14983
Bid Amount	5271480	13058090	43806	130820028	14983

Estimates, Number of Items and Working Days. We find engineer estimates has a significant prediction power on bid amount that is 0.945. Intuitively, contractors submit bids based on the list of items provided by Caltrans and the quantity and unit cost of each item estimated by engineers. The unit price that the contractor can offer should be close to the engineer unit cost estimates. As a consequence, Bid Amount is around Engineer Estimates to some extent due to project size. When control for the number of items in a contract and number of working days, the estimated coefficient on Engineer Estimates does not change much. In the following sections, we will use the variable Engineer estimates to control for contract size based on the discussion above.

Table 18: Regression of Bid Amount on Contract Characteristics

	(1) Bid Amount	(2) Bid Amount	(3) Bid Amount	(4) Bid Amount	(5) Bid Amount
Engineer Estimates	0.945***			0.920***	0.920***
	(0.00261)			(0.00421)	(0.00427)
Number of Items		171981.2***		7347.0***	7213.7***
		(2162.5)		(956.4)	(1110.5)
Working Days			12146.2***		25.59
			(478.5)		(108.3)
_cons	72070.6***	-3333738.5***	453552.3***	-111342.2***	-111545.8***
	(22862.0)	(108120.0)	(144129.9)	(32937.3)	(32952.6)
N	3829	3829	3829	3829	3829
R^2	0.972	0.623	0.144	0.972	0.972
adj. R^2	0.972	0.623	0.144	0.972	0.972

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

4.2 The Model

In this section, we introduce basic settings of our model and discuss on the bidding equilibrium. In Subsection 4.2.1, two assumptions on dependence structure are presented. One requires that the dependency structure among bidders are the same across auctions. The other requires that the dependence structure of private values between any pair of linked bidders is the same, regardless of the auction and bidder characteristics. These assumptions are quite conventional. In Subsection 4.2.2, we show a sufficient condition for the firs-order condition of expected payoff to be held. In the following sections, our reduced-form analysis is mainly based on the idea of the model.

4.2.1 Basic Settings

We consider a first-price, sealed-bid auction model in which $n(\geqslant 2)$ bidders with a given network structure bid for a single indivisible object. Let $G \equiv (N, E)$ characterize the network structure of bidders where $N \equiv \{1, ..., n\}$ collects all bidders and $E \equiv (e_{ij})_{i,j \in N}$ collects all link status between bidders. We consider a symmetric, undirected network structure in which $e_{ij} = e_{ji} = 1$ indicates i and j are linked and $e_{ij} = e_{ji} = 0$ otherwise. There are no self links, $e_{ii} = 0$. Let the neighborhood $N_i \equiv \{j \in N | e_{ij} = 1\}$ collect the bidders that are directly linked to bidder i, and $m_i \equiv |N_i|$ denote the number of neighboring bidders.

Let V_i be bidder i's valuation over the object. Suppose the bidders' value $V = (V_1, ...V_n)$ is drawn from the continuous joint cumulative distribution function $F_V(v)$ with $v = (v_1, ..., v_n)$ denoting realizations of random variables. Let $F_i(v_i)$ denote the marginal cumulative distribution function of bidder i. By Sklar's theorem , there exists a unique copula function C such that

$$F_V(v) = C(F_1(v_1), \cdots, F_n(v_n)). \tag{4.1}$$

Denote X as the vector of auction characteristics.

ASSUMPTION 4.1. *The dependency structure among bidders are the same across auctions, i.e.* $F_{V|X}(v|x) = C(F_{1|X}(v_1|x), \cdots, F_{n|X}(v_n|x))$.

Following Mazo, Girard and Forbes (2015), we consider the Product of Bivariate Copula (PBC) model associated with the network G:

$$C(u_1,...,u_n) = \prod_{\{(i,j):e_{ij}\in E\}} \Gamma_{ij}(u_i^{1/m_i}, u_j^{1/m_j}), \tag{4.2}$$

where Γ_{ij}'s are arbitrary bivariate copulas. For some isolated bidder i, let $\Gamma_i(u_i) = u_i$, so that V_i is independent of all other values. From Mazo, Girard and Forbes (2015), we know that (4.2) is a well defined copula. The PBC model essentially introduces *asymmetry* into the affiliation relationship between bidders, where such asymmetry is fully captured by their network structure G.[4] Specifically, (4.2) implies that the bivariate marginal copula $C_{ij}(u_i, u_j)$ is given by (Proposition 5 in Mazo, Girard and Forbes (2015))

$$C_{ij}(u_i,u_j) = \begin{cases} u_i^{(m_i-1)/m_i} u_j^{(m_j-1)/m_j} \Gamma_{ij}(u_i^{1/m_i}, u_j^{1/m_j}) & \text{if } e_{ij} \in E, \\ u_i u_j & \text{otherwise.} \end{cases} \tag{4.3}$$

Therefore, the values between unlinked bidders are independent, while the value dependency between linked bidders are determined by the copula function Γ_{ij}.

ASSUMPTION 4.2. *(Homogeneous Pair Copula) The dependence structure of private valuations between any pair of linked bidders is the same, regardless of the auction and bidder characteristics. That is,* $\Gamma_{ij}(a_i, a_j) = \Gamma(a_i, a_j; \theta), \forall i, j \in N$.

[4] Moreover, the PBC model is a particular case of the Liebscher copula, which a generic form to construct an asymmetric copula (Liebscher 2008). In fact, PBC is the only Liebscher copula decomposed into bivariate copulas. See Proposition 2 in Mazo, Girard and Forbes (2015).

4.2.2 Equilibrium Bidding

For ease of exposition, assume the joint distribution is symmetric (or exchangeable), i.e. $F_i = F_0, \forall i$. Let $\sigma(\cdot)$ denote the strictly increasing symmetric bidding strategy in equilibrium. Note that here our copula function is no longer symmetric.

Let $W_{-i} \equiv (W_1, ..., W_{i-1}, W_{i+1}, ..., W_n)$ denote the vector of W eliminating W_i. Given the copula model (4.2), the marginal copula of u_{-i} is

$$\begin{aligned} C_{-i}(u_{-i}) &\equiv C(u_1, ..., u_{i-1}, 1, u_{i+1}, ..., u_n) \\ &= \prod_{\{(k,l):k,l \notin N(i), e_{kl} \in E\}} \Gamma(u_k^{1/m_k}, u_l^{1/m_l}) \prod_{\{j:j \in N(i), e_{ij} \in E\}} u_j^{1/m_j}, \end{aligned} \tag{4.4}$$

resulting in the marginal distribution of V_{-i} being

$$\begin{aligned} F_{V_{-i}}(v_{-i}) &= C_{-i}(F_0(v_1), ..., F_0(v_{i-1}), F_0(v_{i+1}), ..., F_0(v_n)) \\ &= \prod_{\{(k,l):k,l \notin N(i), e_{kl} \in E\}} \Gamma(F_0(v_k)^{1/m_k}, F_0(v_l)^{1/m_l}) \prod_{\{j:j \in N(i), e_{ij} \in E\}} F_0(v_j)^{1/m_j}. \end{aligned} \tag{4.5}$$

Bidder i's expected payoff is

$$\begin{aligned} E[\pi(s_i, v_i)] &= (v_i - s_i) P[V_1 \leq \sigma^{-1}(s_i), ..., V_{i-1} \leq \sigma^{-1}(s_i), V_{i+1} \leq v_i, ..., V_n \leq \sigma^{-1}(s_i)] \\ &= (v_i - s_i) F_{V_{-i}}(\sigma^{-1}(s_i), ..., \sigma^{-1}(s_i)) \\ &= (v_i - s_i) C_{-i}(F_0(\sigma^{-1}(s_i)), ..., F_0(\sigma^{-1}(s_i))), \end{aligned} \tag{4.6}$$

where $\sigma(\cdot)$ denotes the symmetric strictly increasing bidding strategy.

The first-order condition yields

$$\sigma'(v_i) = (v_i - \sigma(v_i)) f_0(v_i) \frac{\sum_{j \neq i}^{n} C_{-i}^{(j)}(F_0(v_i), \cdots, F_0(v_i))}{C_{-i}(F_0(v_i), \cdots, F_0(v_i))}, \tag{4.7}$$

where $C_{-i}^{(j)} \equiv \frac{\partial C_{-i}}{\partial u_j}$.

A sufficient condition for equation (4.7) to characterize a unique monotone pure strategy equilibrium (MPSE) is that the copula C satisfies MTP2. It is well known that a function is MTP2 if and only if it is TP2 in all pairs of its arguments (Hubbard, Li and Paarsch 2012). In addition, by Proposition 6 in Mazo, Girard and Forbes (2015), the TP2 property of C_{ij} is inherited from the TP2 property of γ_{ij}.

ASSUMPTION 4.3. *(TP2) The copula* $\gamma(a_i, a_j; \theta)$ *is totally positive of order 2, i.e.* $\Gamma(a_1, a_2; \theta)\Gamma(b_1, b_2; \theta) \geq \Gamma(a_1, b_2; \theta)\Gamma(b_1, a_2; \theta)$*, for all* $a_1 < b_1, a_2 < b_2$*.*

Let G_0 be the marginal distribution of the equilibrium bids, then by GPV,

$$G_0(s_i) = F_0(v_i),\ g_0(s_i) = \frac{f_0(v_i)}{\sigma'(v_i)}. \tag{4.8}$$

Therefore, from the first order condition (4.7) we have

$$v_i = s_i + \frac{C_{-i}(G_0(s_i), \cdots, G_0(s_i))}{g_0(s_i) \sum_{j \neq i} C_{-i}^{(j)}(G_0(s_i), \cdots, G_0(s_i))}. \tag{4.9}$$

ASSUMPTION 4.4. *(Symmetric Pair Copula) The bivariate copula function* $\Gamma(\cdot, \cdot; \theta)$ *is exchangeable, i.e.* $\Gamma(a, b; \theta) = \Gamma(b, a; \theta), \forall a, b \in R$*.*

We derive the expression of $C_{-i}^{(j)}$ in the Appendix.

4.3 Contractor Networks

In this section, we introduce the way we construct contractor networks. The contractor networks are contract-specific. This means, for one project, we construct one network where the nodes are the contractors that submitted bids for the project, and the links between those nodes are defined based on if they share at least one same subcontractor. Detailed construction of networks can be found in Section 4.3.1. In Li and Zhang (2015), they assume private values are symmetrically dependent among all the bidders. However, in our case, we assume the dependency of private values are based on link status. If two bidders are not linked, then their private values are considered independent. In section 4.3.2, we examine the densities of each network to see the extent to which the contractors are connected, hence to what extent their private values are dependent on each other. In subsection 4.3.3, we present the centrality measures that describe contractor importance in the contract-specific networks.

4.3.1 Construction of Networks

For each contract that we observe during 2001 to 2002, we construct a contract-specific bidder network based on all the bidders (contractors) that bid in the same contract. In total, we construct 712 different bidder networks based on 712 different contracts overtime in the sample. The contract-specific bidder network is constructed by the chance of having same subcontractors in the underlying contract. For example, suppose both bidder A and bidder B bid for a contract on date Jun.10. 2001. Bidder A listed a subcontractor, which was also one of bidder B's subcontractors when B bid for the same project, then A and B are considered linked.

Intuitively, we expect that a subcontractor can effectively serve as a go-between that makes information (e.g. costs, qualities, working days, etc.) sharing available for two contractors. According to the regulation of Caltrans bidding process, the contractor must submit its bid along with a list of subcontractors. Naturally, a contractor should have already contacted those subcontractors for their qualities of work and costs for work. For a same project(contract), if two contractors listed at least one same subcontractor, their expectations on unit prices of project items should be similar. As a consequence, the total bid amount submitted by these two contractors should be close too. Detailed reduced-form evidence on dependency of bid amounts based on link status can can be found in Section 4.4.

Figure 12 is an example of how contract-specific networks may look like. Here we plot 36 contract-specific networks from our sample. There are three features about the contract-specific networks. The first feature is that the number of bidders in a contract t, denoted as n_t, can vary. The average number of bidders per contract is 5.05 with a standard deviation of 3.05 (Table 16). The second feature is that some contract-specific networks are complete networks. A complete network is a network where all bidders are connected to each other. For example, the first network in the first row in Figure 12. According to our justification, a complete network indicates that the private values are symmetrically dependent among all contractors. This case coincides with the basic settings in Li and Zhang (2015). The third feature is that there are a small amount of empty networks. An empty network is a network in which there is no links. for example, the fifth network in the first row in Figure 12 is an empty network with three bidders. Notice that there can be two possible cases for a network to be an empty network: there is only one bidder in the network (no self link allowed) or there are more than one bidders but they do not have any same subcontractors. In an empty network, private values are considered independent due to zero links. As long as the contractor network is not a complete network, dependency structures of private values are different from Li and Zhang (2015). In summary, we find that among the 752 networks, 33.78% of them are complete networks and 16.49% of them are empty networks. 66.22% of our contractor networks are not complete networks, indicating a quite different setting compared with the symmetric dependency setting in the current literature.

4.3.2 Densities of Networks

In this subsection, we introduce the densities of contractor networks. Density is a network descriptive measure that equals to the number of links divided by the maximum number of potential links in a network. For example, for a network with n_t bidders, the maximum number of potential links equals to $\frac{n_t*(n_t-1)}{2}$. Density is another straightforward measures that tells us to what extent private values are dependent among a contractor network. If the density of a contractor network equals one, the network is a complete network.

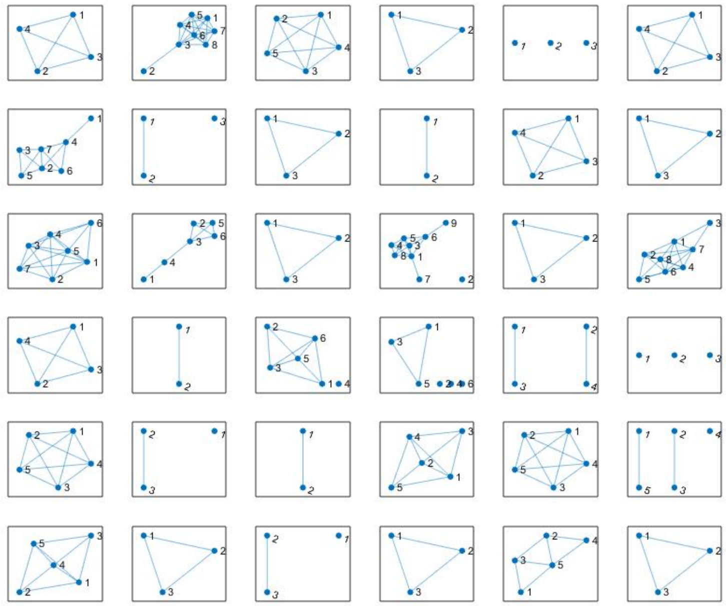


Figure 12: An Example of Network Structure

Figure 13 summarizes the distribution of density of the contractor networks. If we exclude the complete networks and the empty networks, we observe a rather uniform distribution of network densities for very sparse density to very dense density. In our settings, two linked bidders share a dependency structure on their private values, whereas two bidders without a link are considered to have independent private values. In the previous literature, the affiliated private values are assumed to take all pairs of bidders' affiliated values into consideration equally, regardless of network structure. The variation in densities reveals the importance of network structure of the underlying contractors. The traditional way no longer holds because a large fraction of networks are not complete networks and some of them are even very sparse.

4.3.3 Centrality Measures of Networks

In this subsection, we introduce the centrality measures of contractors in its contractor networks. Centrality measures capture a bidder's position/importance in its contract-specific network. We

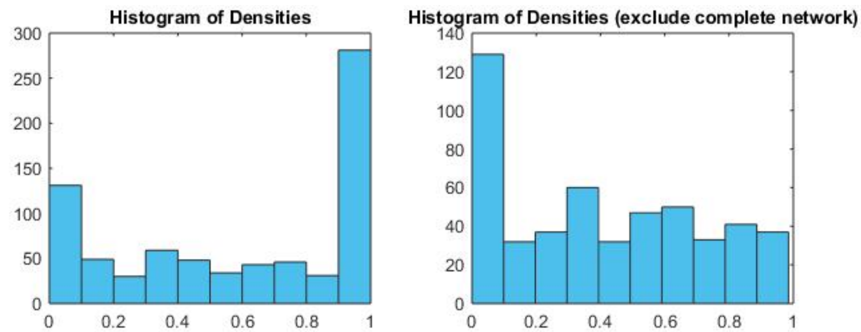


Figure 13: Histogram of Network Densities

use four centrality measures from social network analysis (SNA) to describe a bidder's importance in its network: Degree, Eigenvector, Closeness and Betweenness.

Degree is the simplest and most straightforward centrality measure that counts the number of links directly linked to the bidder. Degree can be interpreted as the immediate risk for a bidder to catch information that flows through a network. Eigenvector is another measure that characterizes the local centrality of a bidder. Different from Degree that treats every direct links equally, Eigenvector highlights the differences in influence level of each links. Closeness is a measure that is defined based on the concept of network paths, whereas Degree and Eigenvector are both defined based on direct links. Closeness tells us the extent to which a bidder is closer to all the other reachable bidders in a network either directly or indirectly. A higher Closeness indicates the bidder is closer to the central of the network. Another measure that is defined based on the concept of network paths is Betweenness. This measure reflects how often an underlying bidder to occur on a randomly chosen shortest path between two randomly chosen bidders in a network.

To summarize, Degree and Eigenvector are good measures for characterizing local importance of the bidder; while Closeness and Betweenness are used for capturing global importance of the bidder. Notice that, for comparability between different networks, we normalize the four centrality measures based on different network size.

In our sample, there are 604 bidders but they repeatedly showed up in 712 different contracts. In the end, we have a total size of 3743 unique contract-bidder observations. For a bidder, it may have different centrality measures in different networks because the network is contract-specific. A detailed summary statistics on centrality measures are given in Table 19. Since we normalize centrality measures for comparison between different networks with various network size, the values of these four centrality measures all lie in the interval from 0 to 1.

Table 19: Summary Statistics: Centrality Measures

Variable	Mean	Std. Dev.	Min.	Max.	N
degree	0.644	0.369	0	1	3743
eigenvector	0.188	0.131	0.014	1	3743
closeness	0.164	0.172	0	1	3743
betweenness	0.024	0.093	0	1	3743

Figure 14 presents distributions of centrality measures on every bidders in each contract. From comparison between the distribution of degree centrality and that of eigenvector centrality, we find that even though many contractors have degree level close to 1, very little of them have eigenvector level around 1. As we introduce, eigenvector accounts for the importance of linked nodes, while degree simply counts linked nodes with equal weights. In Section 4.5, we will see how the importance of a contractor in its network may affect its bidding behavior.

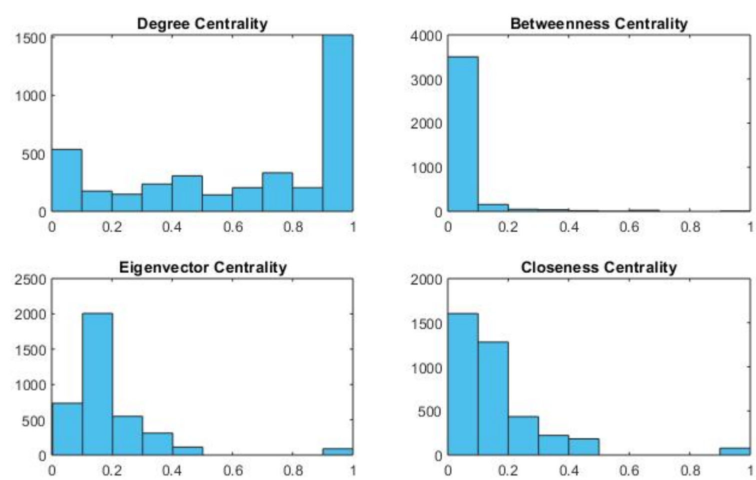


Figure 14: Histogram of Centrality Measures

4.4 Empirical Results: Dependency Structure of Linked Pairs

In this section, we provide key evidence on how private values are dependent based on network structure. In Subsection 4.4.1, we introduce the distance between two contractors. Most pairs of contractors are either unlinked or within two-step reach, that is either directly linked, or indirectly (second-degree) linked. In Subsection 4.4.2 we show that the difference in bidding amount is significantly smaller for two contractors if they are directly linked. This reduced-form finding supports our theory that private values are dependent through network links. We also find that the difference in bid amount are larger for two contractors if they are indirectly linked. In Subsection 4.4.3, we first define a bid gap as the difference in bid amount between the winner and the biggest loser. We then study how the density of contractor network may affect the bid gap. We expect the bid gap to be smaller if the density is higher.

4.4.1 Distance between Contractors

In network studies, distance between two nodes captures the length of the shortest path between two nodes. In an indirect unweighted network, every edge distance is taken to be 1. For example, if node A and node B are directly linked, then the distance between A and B is 1. If node C is directly linked to B but not linked to A, then the distance between A and C is 2. If two nodes are neither directly linked, nor indirectly linked through any node, then they are considered to have infinite distance.

Table 20 summarizes the distances between any two contractors in a specific contractor network in our sample. The total amount of such pairs is 11224, which is calculated by adding up all the maximum potential numbers of links in each contractor network. Around 22.71% of pair of contractors have infinite distance, 63.08% of them are directly linked and over 97.50% of them are within two-step reach or have infinite distance. Since very little fraction of pairs have distance over 2, our reduced-form model in Subsection 4.4.2 will focus on the cases with one-step reach and two-step reach.

Table 20: Distance of Two Contractors

	Count	Frequency
Infinite	2549	22.71
One-step Reach	7080	63.08
Two-step Reach	1330	11.85
Three-step Reach	234	2.08
Four-step Reach	26	0.23
Five-step Reach	5	0.04
Total	11224	100.00

4.4.2 Dependency in Bid Amount between Linked Pairs

Figure 15 plots the distribution of total bid amount of all contractors, divided by engineer estimates in order to control for contract characteristics. We call this ratio between total bid amount and engineer estimates as Bid Amt. If Bid Amt is greater than 1, this means the contractor submits a bid over the amount of engineer estimates; while if Bid Amt is less than 1, this indicates the contractor submits a bid less than the amount of engineer estimates. According to the histogram, we find there are some extreme cases where contractors submit very high bids compared with engineer estimates. There are 2, 3, even 5 times of engineer estimates. We also find that almost half of contractors bid over estimates and half of them bid less than estimates. This variation in Bid Amt reveals variation in bidding behavior of contractors. Next we will study how the linked status may affect bidding behavior between a pair of contractors.

We first construct the dependent variable by calculating the absolute difference between the two Bid Amt of two contractors. For example if Bid Amt of contractor A is 1.3 and Bid Amt of contractor B is 1.1, then the difference in Bid Amt is 0.2. We collect all the difference in Bid Amt for all the pairs of contractors in each contractor network. In the first column, our independent variable is an indicator variable that equals one indicating two contractors are directly linked and zero

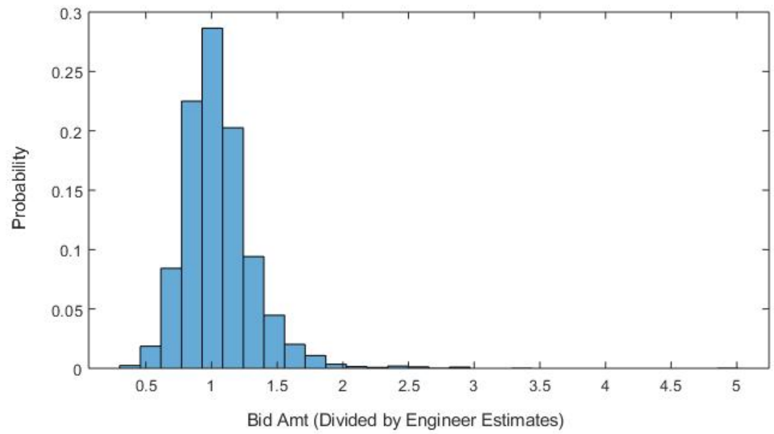


Figure 15: Probability Density Distribution of Bid Amount

otherwise. In the second column, our independent variable is an indicator variable that equals one indicating two contractors are linked in the second-degree (indirectly linked) and zero otherwise.

Table 21 presents the linear regression results of difference in bidding amount on linked status. We find the effect of directly linked status on difference in Bid Amt is significantly negative with magnitude 0.0681 at 99% confidence level. This indicates that if two contractors are linked in its contractor network when bidding for the same contract, then they are likely to bid a similar price compared with unlinked or indirectly linked. Intuitively, if two contractors share a same subcontractor when they bid for the same contract, they are likely to get similar price or qualities from the same subcontractor. The subcontractor then serves as a go-between that makes information between two contractors shareable.

We also run a specification on pairs that are second-degree linked. We find the effect of second-degree linked status on difference in Bid Amt is insignificant. This indicates that indirect links have zero effect on bid amount. Intuitively, second-degree link involves two different subcontractors that brings possibly quite different information on bids. As a consequence, the two contractors that are indirectly linked will also have quite various bid behaviour.

Table 21: Regression of Difference in Bidding Amount on Linked Status

	(1) Difference in Bid Amt	(2) Difference in Bid Amt
Directly Linked	-0.0681***	
	(0.00397)	
Second-degree Linked		0.00571
		(0.00594)
Constant	0.228***	0.183***
	(0.00318)	(0.00206)
N	11047	11047
R^2	0.026	0.000
adj. R^2	0.026	0.000

Standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

4.4.3 Bid Gap and Network Densities

In this subsection, we define the bid gap for a bid and investigate how network densities may affect bid gaps. We expect that if a contractor network has higher density, then contractors are engaged more often by sharing same subcontractors. As a consequence, when they submit bids, they tend to submit quite close prices over the project.

Bid gap is defined as the difference on total bid amount between the winner and the biggest loser (the contractor that submit the highest bid), divided by engineer estimates. Besides, we also control for the number of contractors in each contract to make such gaps comparable. The dependent variable in Column (1) in Table 22 is bid gap controlling for engineer estimates. The dependent variable in Column (2) is the bid gap from (1) divided by the number of contractors in the associated contractor network. Dependent variable in Column (2) can be interpreted as the unit bid gap (bid gap per contractor) for a contract.

From Table 22, we find the effect of density on bid gap for both cases are significantly negative at level 99%. Negative effect coincides with our initial expectation that if a contractor network is more dense, then the bid gap between the winner and the biggest loser is smaller. In other words, if there are more links in a network, then the contractors tend to submit similar bids compared with the case with less links. In the case where network density changes from 0 to 1, equivalently the network changes from an empty network to a complete network, the bid gap decreases by 0.209 significantly and the unit bid gap decreases by 0.0339 significantly.

Table 22: Regression of Bid Gap on Density

	(1)	(2)
	Bid Gap	Bid Gap (control for the number of contractors)
density	-0.209***	-0.0339***
	(0.0360)	(0.00718)
_cons	0.467***	0.0860***
	(0.0270)	(0.00539)
N	677	677
R^2	0.047	0.032
adj. R^2	0.046	0.031

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

4.5 Empirical Results: Network Effects

In this section, we examine how the position of a contractor in its contractor network may affect its bidding behavior and probability of winning. We first run a linear regression of total bid amount (divided by engineer estimates) on the four centrality measures. Then we conduct a Probit model of probability of winning on centrality measures. In the end, we find, the importance or the position of a contractor in its contractor network does have certain effects on its bidding behavior and probability of winning.

Table 23 presents regression results of Bid Amt (total bid amount divided by engineer estimates) on four centrality measures. In the Caltrans auction, the contractor with the lowest total bid amount will win the contract. From Column (6) in Table 23, we find that all the four centrality measures have significant effect on Bid Amt, but the signs of the coefficients have different directions. Recall that both degree centrality and eigenvector centrality capture the local importance of a node, while closeness centrality and betweenness centrality capture the global importance of a node. Column (6) in Table 23 indicates that if a contractor is locally important, then it tends to bid with smaller price; if a contractor is globally important, then it tends to bid with higher price. Intuitively, local importance is crucial to our case because we have examined that direct linked status have significant effect on bidding behaviors of a pair of contractors.

Table 23: Linear Regression: Bid-Engineer Estimates Ratio on Centrality Measures

	(1)	(2)	(3)	(4)	(5)	(6)
	Bid Amt	Bid Amt	Bid Amt	Bid Amt	Bid Amt	Bid Amt
degree	-0.0782***				-0.149***	-0.156***
	(0.0121)				(0.0162)	(0.0163)
eigenvector		-0.00298			-0.126***	-0.133***
		(0.0290)			(0.0327)	(0.0327)
closeness			0.0138		0.237***	0.243***
			(0.0262)		(0.0363)	(0.0363)
betweenness				0.105**		0.154***
				(0.0451)		(0.0449)
_cons	1.102***	1.052***	1.050***	1.049***	1.135***	1.135***
	(0.00904)	(0.00728)	(0.00621)	(0.00467)	(0.0119)	(0.0119)
N	3801	3801	3801	3801	3801	3801
R^2	0.011	0.000	0.000	0.001	0.022	0.025
adj. R^2	0.011	0.000	0.000	0.001	0.021	0.024

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Next, we conduct a reduced-form analysis on the effect of network centrality measures on Win. Variable Win is a binary variable with $Win = 1$ represents the contractor win the auction; $Win = 0$ otherwise. We run a Probit model of Win on Degree, Eigenvector, Closeness and Betweenness centrality measures. We are interested in whether the position of a bidder in network has significant effect on whether the bidder win the auction. Table 24 presents the estimation results.

Degree centrality has a significantly negative effect on winning probability. An explanation would be that with more neighbors that share the same subcontractors, more contractors bid for less values, in turn reducing the likelihood of winning. However, eigenvector centrality has a significantly positive effect on winning probability. Due to higher importance of linked contractors, the contractor itself can gain more valuable information on the cost of projects, hence more likely to win the bid. Closeness centrality also has a significant positive effect on probability of winning. Higher closeness means the node is more central in the whole network. A contractor that lies in the central of the network can effectively gain useful information and as a consequence more likely to win a bid.

Table 24: Probit Model: Win on Centrality Measures

	(1) win	(2) win	(3) win	(4) win	(5) win	(6) win
degree	0.0645				-0.257***	-0.255***
	(0.0646)				(0.0883)	(0.0888)
eigenvector		1.689***			1.083***	1.087***
		(0.164)			(0.197)	(0.198)
closeness			1.167***		1.128***	1.126***
			(0.127)		(0.186)	(0.186)
betweenness				0.0601		-0.0629
				(0.250)		(0.248)
_cons	-0.954***	-1.252***	-1.119***	-0.914***	-1.164***	-1.164***
	(0.0483)	(0.0416)	(0.0334)	(0.0247)	(0.0664)	(0.0664)
N	3743	3743	3743	3743	3743	3743

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

4.6 Conclusion

Auction with affiliated valuations are widely studied. Previous literature on affiliation require that the dependency structures of every pair of bidders are the same. In this paper, we show some empirical evidence on how link status in a bidder network matters when studying affiliated values. We can draw two main conclusions based on our empirical findings. First, we find if two contractors are linked in its contractor network, then they are significantly likely to submit close bid prices. Second, the position of a contractor in its contractor network has some significant effects on its bidding amount and probability of winning. There could be two feasible and enlightening tasks to be considered as part of our future work. One is a structural study on our model with estimation methods. The other could be a new work that we can relax the assumption that requires all linked pairs to have same dependent structures, by taking the link weights into consideration.

5 CONCLUSION

Network effect has been widely studied in Econometrics in various of aspects. Researchers are interested in how the network structure may affect certain network outcomes. In this dissertation, we studied different topics of network effect studies. In the first essay, we studied firm networks and underwriter networks to see how they may affect the public firms' post IPO performance. In the second essay, we investigated how to identify and estimate peer effects in latent networks. In the third essay, we used a network perspective to analyze the affiliated bidding behavior in first-price auction. From the works, we find that network structure does have significant effects on some outcomes we care about. Future studies can be done around studies on network formation and network effects. We can relax the networks to be endogenous and see what would the network effect be.

REFERENCES

Ahern, Kenneth. 2013. "Network centrality and the cross section of stock returns. SSRN Scholarly Paper ID 2197370." *Social Science Research Network, Rochester, NY* .

Ahern, Kenneth R and Jarrad Harford. 2014. "The importance of industry links in merger waves." *The Journal of Finance* 69(2):527–576.

Andrew, Alan L, K-W Eric Chu and Peter Lancaster. 1993. "Derivatives of eigenvalues and eigenvectors of matrix functions." *SIAM Journal on Matrix Analysis and Applications* 14(4):903–926.

Andrews, Donald WK. 1994. "Asymptotics for semiparametric econometric models via stochastic equicontinuity." *Econometrica* 62(1):43–72.

Ardichvili, Alexander, Richard Cardozo and Sourav Ray. 2003. "A theory of entrepreneurial opportunity identification and development." *Journal of Business Venturing* 18(1):105–123.

Bajo, Emanuele, Thomas J Chemmanur, Karen Simonyan and Hassan Tehranian. 2016. "Underwriter networks, investor attention, and initial public offerings." *Journal of Financial Economics* 122(2):376–408.

Bengtsson, Maria and Sören Kock. 2000. "Coopetition in business Networks to cooperate and compete simultaneously." *Industrial Marketing Management* 29(5):411–426.

Blume, Lawrence E, William A Brock, Steven N Durlauf and Rajshri Jayaraman. 2015. "Linear social interactions models." *Journal of Political Economy* 123(2):444–496.

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

Burt, Ronald S. 2000. "The network structure of social capital." *Research in Organizational Behavior* 22:345–423.

Calvó-Armengol, Antoni, Eleonora Patacchini and Yves Zenou. 2009. "Peer effects and social networks in education." *The Review of Economic Studies* 76(4):1239–1267.

Case, Anne C. 1991. "Spatial patterns in household demand." *Econometrica* 59(4):953–965.

Chandrasekhar, Arun and Randall Lewis. 2011. "Econometrics of sampled networks." *Unpublished Manuscript, MIT.[422]* .

Chu, Yongqiang. 2012. "Optimal capital structure, bargaining, and the supplier market structure." *Journal of Financial Economics* 106(2):411–426.

Cohen, Lauren and Andrea Frazzini. 2008. "Economic links and predictable returns." *The Journal of Finance* 63(4):1977–2011.

Cooper, Arnold C. 2017. "Networks, alliances, and entrepreneurship." *Strategic Entrepreneurship: Creating a New Mindset* pp. 201–222.

De Paula, Áureo, Seth Richards-Shubik and Elie Tamer. 2018. "Identifying preferences in networks with bounded degree." *Econometrica* 86(1):263–288.

Deeds, David L and Charles WL Hill. 1996. "Strategic alliances and the rate of new product development: An empirical study of entrepreneurial biotechnology firms." *Journal of Business Venturing* 11(1):41–55.

Dietz, Thomas. 2012. "Contract Law, Relational Contracts, and Reputational Networks in International Trade: An Empirical Investigation into Cross-Border Contracts in the Software Industry." *Law & Social Inquiry* 37(1):25–57.

Dubini, Paola and Howard Aldrich. 1991. "Personal and extended networks are central to the entrepreneurial process." *Journal of Business Venturing* 6(5):305–313.

El-Khatib, Rwan, Kathy Fogel and Tomas Jandik. 2015. "CEO network centrality and merger performance." *Journal of Financial Economics* 116(2):349–382.

Gargiulo, Martin and Mario Benassi. 2000. "Trapped in your own net? Network cohesion, structural holes, and the adaptation of social capital." *Organization Science* 11(2):183–196.

Genest, Christian, Kilani Ghoudi and L-P Rivest. 1995. "A semiparametric estimation procedure of dependence parameters in multivariate families of distributions." *Biometrika* 82(3):543–552.

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

Granovetter, Mark. 1985. "Economic action and social structure: The problem of embeddedness." *American Journal of Sociology* 91(3):481–510.

Gulati, Ranjay. 1999. "Network location and learning: The influence of network resources and firm capabilities on alliance formation." *Strategic Management Journal* 20(5):397–420.

Hanna, Victoria and Kathryn Walsh. 2002. "Small firm networks: a successful approach to innovation?" *R&D Management* 32(3):201–207.

Hennessy, Christopher A and Dmitry Livdan. 2009. "Debt, bargaining, and credibility in firm–supplier relationships." *Journal of Financial Economics* 93(3):382–399.

Hirshleifer, David and Siew Hong Teoh. 2003. "Limited attention, information disclosure, and financial reporting." *Journal of Accounting and Economics* 36(1-3):337–386.

Hong, Harrison and Jeremy C Stein. 1999. "A unified theory of underreaction, momentum trading, and overreaction in asset markets." *The Journal of Finance* 54(6):2143–2184.

Hu, Yingyao. 2008. "Identification and estimation of nonlinear models with misclassification error using instrumental variables: A general solution." *Journal of Econometrics* 144(1):27–61.

Hu, Yingyao. 2017. "The econometrics of unobservables: Applications of measurement error models in empirical industrial organization and labor economics." *Journal of Econometrics* 200(2):154–168.

Hu, Yingyao and Susanne M Schennach. 2008. "Instrumental variable treatment of nonclassical measurement error models." *Econometrica* 76(1):195–216.

Hubbard, Timothy P, Tong Li and Harry J Paarsch. 2012. "Semiparametric estimation in models of first-price, sealed-bid auctions with affiliation." *Journal of Econometrics* 168(1):4–16.

Huberman, Gur and Tomer Regev. 2001. "Contagious speculation and a cure for cancer: A non-event that made stock prices soar." *The Journal of Finance* 56(1):387–396.

Kelley, Donna J and Mark P Rice. 2002. "Advantage beyond founding: The strategic use of technologies." *Journal of Business Venturing* 17(1):41–57.

Leary, Mark T and Michael R Roberts. 2014. "Do peer firms affect corporate financial policy?" *The Journal of Finance* 69(1):139–178.

Lechner, Christian, Michael Dowling and Isabell Welpe. 2006. "Firm networks and firm development: The role of the relational mix." *Journal of Business Venturing* 21(4):514–540.

Li, Tong and Bingyu Zhang. 2015. "Affiliation and entry in first-price auctions with heterogeneous bidders: An analysis of merger effects." *American Economic Journal: Microeconomics* 7(2):188–214.

Li, Tong, Isabelle Perrigne and Quang Vuong. 2002. "Structural estimation of the affiliated private value auction model." *RAND Journal of Economics* pp. 171–193.

Liebscher, Eckhard. 2008. "Construction of asymmetric multivariate copulas." *Journal of Multivariate Analysis* 99(10):2234–2250.

Lipparini, Andrea and Maurizio Sobrero. 1997. "Co-ordinating multi-firm innovative processes: entrepreneur as catalyst in small-firm networks." *The Formation of Inter-organizational Networks* pp. 199–219.

Liu, Laura Xiaolei, Ann E Sherman and Yong Zhang. 2014. "The long-run role of the media: Evidence from initial public offerings." *Management Science* 60(8):1945–1964.

Mahajan, Aprajit. 2006. "Identification and estimation of regression models with misclassification." *Econometrica* 74(3):631–665.

Mazo, Gildas, Stephane Girard and Florence Forbes. 2015. "A class of multivariate copulas based on products of bivariate copulas." *Journal of Multivariate Analysis* 140:363–376.

Milgrom, Paul R and Robert J Weber. 1982. "A theory of auctions and competitive bidding." *Econometrica: Journal of the Econometric Society* pp. 1089–1122.

Newey, Whitney K. 1994. "The asymptotic variance of semiparametric estimators." *Econometrica* 62(6):1349–1382.

Newey, Whitney K and Daniel McFadden. 1994. "Large sample estimation and hypothesis testing." *Handbook of Econometrics* 4:2111–2245.

Newey, Whitney K and James L Powell. 2003. "Instrumental variable estimation of nonparametric models." *Econometrica* 71(5):1565–1578.

Pappas, James M and Bill Wooldridge. 2007. "Middle managers' divergent strategic activity: An investigation of multiple measures of network centrality." *Journal of Management Studies* 44(3):323–341.

Peng, Lin and Wei Xiong. 2006. "Investor attention, overconfidence and category learning." *Journal of Financial Economics* 80(3):563–602.

Rio, Emmanuel. 2017. *Asymptotic theory of weakly dependent random processes*. Springer.

Snehota, Ivan and Hakan Hakansson. 1995. *Developing relationships in business networks*. Routledge London.

Uzzi, Brian. 1997. "Social structure and competition in interfirm networks: The paradox of embeddedness." *Administrative Science Quarterly* pp. 35–67.

White, Halbert. 1996. *Estimation, inference and specification analysis*. Cambridge university press.

Wooldridge, Jeffrey M. 1994. "Estimation and inference for dependent processes." *Handbook of Econometrics* 4:2639–2738.

APPENDIX A

In Appendix A, we will show how the full rank conditions (Assumption 3.6) be easily satisfied using a simple example. Note that we can always emprically test the full rank condition. Consider a case where $X_i \in \{0,1\}$ is a one-dimensional, strictly exogenous, binary characteristic. For example, in the study of social network, such characteristic can be gender; in the study of firm network, such can be whether a firm has *R&D* investments. As introduced, we are able to observe two measurements over two networks, peer-effects networks A and A', and contextual-effects networks C and C', while the latent networks are A^* and C^* correspondingly. Define A_i^* and C_i^* as the neighborhood of individual i in networks A^* and C^*, respectively. So that the cardinalities $|A_i^*|$ and $|C_i^*|$ represent the number of individuals that are directly linked to individual i in the two networks, respectively. Similar definitions for observed networks always hold.

As we introduced, $Z_{i1}^* \equiv \sum_{j=1}^n C_{ij}^* X_j$ and $Z_{i2}^* \equiv \sum_{j=1}^n \sum_{k=1}^n A_{ij}^* C_{jk}^* X_k$. Under this simple case, we can rewrite Z_{i1}^* and Z_{i2}^* as $Z_{i1}^* = \frac{1}{|C_i^*|} \sum_{j \in C_i^*} X_j$ and $Z_{i2}^* = \frac{1}{\sum_{k \in A_i^*} |C_k^*|} \sum_{j \in C_k^*} X_j$, respectively (and similarly for $Z_{i1}, Z_{i2}, Z'_{i1}, Z'_{i2}$). Suppose for simplicity, the probabilities of observing the right link status between two individuals i and j in network A and C are given by $Pr(A_{ij} = 1|A_{ij}^* = 1) = Pr(A_{ij} = 0|A_{ij}^* = 0) = \lambda$ and $Pr(C_{ij} = 1|C_{ij}^* = 1) = Pr(C_{ij} = 0|C_{ij}^* = 0) = r$, respectively. Note that observing the right link status between two individuals includes two possible cases: one is observing a link when two individuals are actually linked and the other is not observing a link when two individuals are not actually linked. Suppose in the population, the fraction of male is given by q.

We start with figuring out the probability matrices for Z_1 conditional on Z_1^* which only contains linking information in network C (similar for Z'_1 conditional on Z_1^*). Overall there are four cases to discuss if we impose the relaxed version of Assumption 3.5 that $|C_i| \geq |C_i^*|$ and $|C'_i| \geq |C_i^*|$, for $\forall i \in N$. For simplicity, we start with the case where $|C_i^*| = 2$ and then extend to general case where

$|C_i^*| = n$.

Case 1:$|C_i| = |C_i'| = |C_i^*| = 2$

$$M_{Z_1|Z_1^*} = \begin{bmatrix} Pr(Z_1=0|Z_1^*=0) & Pr(Z_1=0|Z_1^*=\frac{1}{2}) & Pr(Z_1=0|Z_1^*=1) \\ Pr(Z_1=\frac{1}{2}|Z_1^*=0) & Pr(Z_1=\frac{1}{2}|Z_1^*=\frac{1}{2}) & Pr(Z_1=\frac{1}{2}|Z_1^*=1) \\ Pr(Z_1=1|Z_1^*=0) & Pr(Z_1=1|Z_1^*=\frac{1}{2}) & Pr(Z_1=1|Z_1^*=1) \end{bmatrix}_{3\times 3}$$

$$M_{Z_1'|Z_1^*} = \begin{bmatrix} Pr(Z_1'=0|Z_1^*=0) & Pr(Z_1'=0|Z_1^*=\frac{1}{2}) & Pr(Z_1'=0|Z_1^*=1) \\ Pr(Z_1'=\frac{1}{2}|Z_1^*=0) & Pr(Z_1'=\frac{1}{2}|Z_1^*=\frac{1}{2}) & Pr(Z_1'=\frac{1}{2}|Z_1^*=1) \\ Pr(Z_1'=1|Z_1^*=0) & Pr(Z_1'=1|Z_1^*=\frac{1}{2}) & Pr(Z_1'=1|Z_1^*=1) \end{bmatrix}_{3\times 3}$$

$$M_{Z_1,Z_1'} = \begin{bmatrix} Pr(Z_1=0,Z_1'=0) & Pr(Z_1=0,Z_1'=\frac{1}{2}) & Pr(Z_1=0,Z_1'=1) \\ Pr(Z_1=\frac{1}{2},Z_1'=0) & Pr(Z_1=\frac{1}{2},Z_1'=\frac{1}{2}) & Pr(Z_1=\frac{1}{2},Z_1'=1) \\ Pr(Z_1=1,Z_1'=0) & Pr(Z_1=1,Z_1'=\frac{1}{2}) & Pr(Z_1=1,Z_1'=1) \end{bmatrix}_{3\times 3}$$

$$Pr(Z_1=0|Z_1^*=0) = r^2 + C_2^1 r(1-r)q + (1-r)^2 q^2$$

$$Pr(Z_1=\tfrac{1}{2}|Z_1^*=0) = C_2^1 r(1-r)(1-q) + C_2^1 (1-r)^2 q(1-q)$$

$$Pr(Z_1=1|Z_1^*=0) = (1-r)^2(1-q)^2$$

$$Pr(Z_1=0|Z_1^*=\tfrac{1}{2}) = r(1-r)q + (1-r)^2 q^2$$

$$Pr(Z_1=\tfrac{1}{2}|Z_1^*=\tfrac{1}{2}) = r^2 + r(1-r) + C_2^1 (1-r)^2 q(1-q)$$

$$Pr(Z_1=1|Z_1^*=\tfrac{1}{2}) = r(1-r)(1-q) + (1-r)^2(1-q)^2$$

$$Pr(Z_1=0|Z_1^*=1) = (1-r)^2 q^2$$

$$Pr(Z_1=\tfrac{1}{2}|Z_1^*=1) = C_2^1 r(1-r)q + C_2^1 (1-r)^2 q(1-q)$$

$$Pr(Z_1=1|Z_1^*=1) = r^2 + C_2^1 r(1-r)(1-q) + (1-r)^2(1-q)^2$$

Case 2:$|C_i| = 3 > |C_i'| = |C_i^*| = 2$

$$M_{Z_1|Z_1^*} = \begin{bmatrix} Pr(Z_1=0|Z_1^*=0) & Pr(Z_1=0|Z_1^*=\frac{1}{2}) & Pr(Z_1=0|Z_1^*=1) \\ Pr(Z_1=\frac{1}{3}|Z_1^*=0) & Pr(Z_1=\frac{1}{3}|Z_1^*=\frac{1}{2}) & Pr(Z_1=\frac{1}{3}|Z_1^*=1) \\ Pr(Z_1=\frac{2}{3}|Z_1^*=0) & Pr(Z_1=\frac{2}{3}|Z_1^*=\frac{1}{2}) & Pr(Z_1=\frac{2}{3}|Z_1^*=1) \\ Pr(Z_1=1Z_1^*=0) & Pr(Z_1=1|Z_1^*=\frac{1}{2}) & Pr(Z_1=1|Z_1^*=1) \end{bmatrix}_{4\times 3}$$

$$M_{Z_1'|Z_1^*} = \begin{bmatrix} Pr(Z_1'=0|Z_1^*=0) & Pr(Z_1'=0|Z_1^*=\frac{1}{2}) & Pr(Z_1'=0|Z_1^*=1) \\ Pr(Z_1'=\frac{1}{2}|Z_1^*=0) & Pr(Z_1'=\frac{1}{2}|Z_1^*=\frac{1}{2}) & Pr(Z_1'=\frac{1}{2}|Z_1^*=1) \\ Pr(Z_1'=1|Z_1^*=0) & Pr(Z_1'=1|Z_1^*=\frac{1}{2}) & Pr(Z_1'=1|Z_1^*=1) \end{bmatrix}_{3\times 3}$$

$$M_{Z_1,Z_1'} = \begin{bmatrix} Pr(Z_1=0,Z_1'=0) & Pr(Z_1=0,Z_1'=\frac{1}{2}) & Pr(Z_1=0,Z_1'=1) \\ Pr(Z_1=\frac{1}{3},Z_1'=0) & Pr(Z_1=\frac{1}{3},Z_1'=\frac{1}{2}) & Pr(Z_1=\frac{1}{3},Z_1'=1) \\ Pr(Z_1=\frac{2}{3},Z_1'=0) & Pr(Z_1=\frac{2}{3},Z_1'=\frac{1}{2}) & Pr(Z_1=\frac{2}{3},Z_1'=1) \\ Pr(Z_1=1,Z_1'=0) & Pr(Z_1=1,Z_1'=\frac{1}{2}) & Pr(Z_1=1,Z_1'=1) \end{bmatrix}_{4\times 3}$$

$$Pr(Z_1=0|Z_1^*=0) = r^2q + C_2^1 r(1-r)q^2 + (1-r)^2q^3$$

$$Pr(Z_1=\tfrac{1}{3}|Z_1^*=0) = r^2(1-q) + C_2^1C_2^1 r(1-r)q(1-q) + C_3^1(1-r)^2q^2(1-q)$$

$$Pr(Z_1=\tfrac{2}{3}|Z_1^*=0) = C_2^1 r(1-r)(1-q)^2 + C_3^1(1-r)^2q(1-q)^2$$

$$Pr(Z_1=1|Z_1^*=0) = (1-r)^2(1-q)^3$$

$$Pr(Z_1=0|Z_1^*=\tfrac{1}{2}) = r(1-r)q^2 + (1-r)^2q^3$$

$$Pr(z_1=\tfrac{1}{3}|Z_1^*=\tfrac{1}{2}) = r^2q + C_2^1 r(1-r)q(1-q) + r(1-r)q^2 + C_3^1(1-r)^2q^2(1-q)$$

$$Pr(Z_1=\tfrac{2}{3}|Z_1^*=\tfrac{1}{2}) = r^2(1-q) + r(1-r)(1-q)^2 + C_2^1 r(1-r)q(1-q) + C_3^1(1-r)^2q(1-q)^2$$

$$Pr(Z_1=1|Z_1^*=\tfrac{1}{2}) = r(1-r)(1-q)^2 + (1-r)^2(1-q)^3$$

$$Pr(Z_1=0|Z_1^*=1) = (1-r)^2q^3$$

$$Pr(Z_1=\tfrac{1}{3}|Z_1^*=1) = C_2^1 r(1-r)q^2 + C_3^1(1-r)^2q^2(1-q)$$

$$Pr(Z_1=\tfrac{2}{3}|Z_1^*=1) = r^2q + C_2^1C_2^1 r(1-r)q(1-q) + C_3^1(1-r)^2q(1-q)^2$$

$$Pr(Z_1=1|Z_1^*=1) = r^2(1-q) + C_2^1 r(1-r)(1-q)^2 + (1-r)^2(1-q)^3$$

Case 3:$|C_i| = 3 = |C_i'| = 3 > |C_i^*| = 2$

$$M_{Z_1|Z_1^*} = \begin{bmatrix} Pr(Z_1 = 0|Z_1^* = 0) & Pr(Z_1 = 0|Z_1^* = \frac{1}{2}) & Pr(Z_1 = 0|Z_1^* = 1) \\ Pr(Z_1 = \frac{1}{3}|Z_1^* = 0) & Pr(Z_1 = \frac{1}{3}|Z_1^* = \frac{1}{2}) & Pr(Z_1 = \frac{1}{3}Z_1^* = 1) \\ Pr(Z_1 = \frac{2}{3}|Z_1^* = 0) & Pr(Z_1 = \frac{2}{3}|Z_1^* = \frac{1}{2}) & Pr(Z_1 = \frac{2}{3}|Z_1^* = 1) \\ Pr(Z_1 = 1|Z_1^* = 0) & Pr(Z_1 = 1|Z_1^* = \frac{1}{2}) & Pr(Z_1 = 1|Z_1^* = 1) \end{bmatrix}_{4\times 3}$$

$$M_{Z_1'|Z_1^*} = \begin{bmatrix} Pr(Z_1' = 0|Z_1^* = 0) & Pr(Z_1' = 0|Z_1^* = \frac{1}{2}) & Pr(Z_1' = 0|Z_1^* = 1) \\ Pr(Z_1' = \frac{1}{3}|Z_1^* = 0) & Pr(Z_1' = \frac{1}{3}|Z_1^* = \frac{1}{2}) & Pr(Z_1' = \frac{1}{3}|Z_1^* = 1) \\ Pr(Z_1' = \frac{2}{3}|Z_1^* = 0) & Pr(Z_1' = \frac{2}{3}|Z_1^* = \frac{1}{2}) & Pr(Z_1' = \frac{2}{3}|Z_1^* = 1) \\ Pr(Z_1' = 1|Z_1^* = 0) & Pr(Z_1' = 1|Z_1^* = \frac{1}{2}) & Pr(Z_1' = 1|Z_1^* = 1) \end{bmatrix}_{4\times 3}$$

Case 4:$|C_i| = 4 > |C_i'| = 3 > |C_i^*| = 2$

$$M_{Z_1|Z_1^*} = \begin{bmatrix} Pr(Z_1 = 0|Z_1^* = 0) & Pr(Z_1 = 0|Z_1^* = \frac{1}{2}) & Pr(Z_1 = 0|Z_1^* = 1) \\ Pr(Z_1 = \frac{1}{4}|Z_1^* = 0) & Pr(Z_1 = \frac{1}{4}|Z_1^* = \frac{1}{2}) & Pr(Z_1 = \frac{1}{4}|Z_1^* = 1) \\ Pr(Z_1 = \frac{1}{2}|Z_1^* = 0) & Pr(Z1 = \frac{1}{2}|Z_1^* = \frac{1}{2}) & Pr(Z_1 = \frac{1}{2}|Z_1^* = 1) \\ Pr(Z_1 = \frac{3}{4}|Z_1^* = 0) & Pr(Z_1 = \frac{3}{4}|Z_1^* = \frac{1}{2}) & Pr(Z_1 = \frac{3}{4}|Z_1^* = 1) \\ Pr(Z_1 = 1|Z_1^* = 0) & Pr(Z_1 = 1|Z_1^* = \frac{1}{2}) & Pr(Z_1 = 1|Z_1^* = 1) \end{bmatrix}_{5\times 3}$$

$$M_{Z_1'|Z_1^*} = \begin{bmatrix} Pr(Z_1' = 0|Z_1^* = 0) & Pr(Z_1' = 0|Z_1^* = \frac{1}{2}) & Pr(Z_1' = 0|Z_1^* = 1) \\ Pr(Z_1' = \frac{1}{3}|Z_1^* = 0) & Pr(Z_1' = \frac{1}{3}|Z_1^* = \frac{1}{2}) & Pr(Z_1' = \frac{1}{3}|Z_1^* = 1) \\ Pr(Z_1' = \frac{2}{3}|Z_1^* = 0) & Pr(Z_1' = \frac{2}{3}|Z_1^* = \frac{1}{2}) & Pr(Z_1' = \frac{2}{3}|Z_1^* = 1) \\ Pr(Z_1' = 1|Z_1^* = 0) & Pr(Z_1' = 1|Z_1^* * = \frac{1}{2}) & Pr(Z_1' = 1|Z_1^* = 1) \end{bmatrix}_{4\times 3}$$

We can always extend the previous example to a more general case where $|C_i| = m$, $|C_i^*| = 2$, and $m \geq 2$. Suppose $|X_j = 1| = k$, that is we observe k female links in the neighborhood of individual i in total.

$$M_{Z_1|Z_1^*} = \begin{bmatrix} \\ Pr(Z_1 = \frac{k}{m}|Z_1^* = 0) & Pr(Z_1 = \frac{k}{m}|Z_1^* = \frac{1}{2}) & Pr(Z_1 = \frac{k}{m}|Z_1^* = 1) \\ \\ \end{bmatrix}_{(m+1)\times 3},$$

where

$$\begin{aligned} Pr(Z_1 = \frac{k}{m}|Z_1^* = 0) =& r^2 C_{m-2}^k q^{m-2-k}(1-q)^k + C_2^1 r(1-r) C_{m-1}^k q^{m-1-k}(1-q)^k + \\ & (1-r)^2 C_m^k q^{m-k}(1-q)^k \end{aligned}$$

$$\begin{aligned} Pr(Z_1 = \frac{k}{m}|Z_1^* = \frac{1}{2}) =& r^2 C_{m-2}^{k-1} q^{m-1-k}(1-q)^{k-1} + r(1-r) \\ & [C_{m-1}^k q^{m-1-k}(1-q)^k + C_{m-1}^{k-1} q^{m-k}(1-q)^{k-1}] + \\ & (1-r)^2 C_m^k q^{m-k}(1-q)^k \end{aligned}$$

$$\begin{aligned} Pr(Z_1 = \frac{k}{m}|Z_1^* = 1) =& r^2 C_{m-2}^{k-2} q^{m-k}(1-q)^{k-2} + C_2^1 r(1-r) C_{m-1}^{k-1} q^{m-k}(1-q)^{k-1} \\ & + (1-r)^2 C_m^k q^{m-k}(1-q)^k \end{aligned}$$

Furthermore, we can extend $|C_i^*| = 2$ to $|C_i^*| = n$. Keep $|C_i| = m$ and $m \geq n$. Suppose $|X_j = 1| = k$ and $|X_j^* = 1| = k^*$, that is we observe k female links in total, while the true number of female links is k^*.

$$M_{Z_1|Z_1^*} = \begin{bmatrix} \\ Pr(Z_1 = \frac{k}{m}|Z_1^* = \frac{k^*}{n}) \\ \\ \end{bmatrix}_{(m+1)\times(n+1)},$$

where

$$Pr(Z_1 = \frac{k}{m}|Z_1^* = \frac{k^*}{n}) = \sum_{w=0}^{n} r^{n-w}(1-r)^w \left[\sum_{s=0}^{min\{n-w,k^*\}} C_{m-(n-w)}^{k-s} q^{m-(n-w)-(k-s)}(1-q)^{k-s}\right]$$

k/k^*: number of observed/true female links.

m/n: number of observed/true links.

w: possible number of links mis-measured among true links ($0 \leq w \leq n$).

s: number of female links correctly measured.

$m-(n-w)$: total number of links mis-measured.

Next, we move on to figure out $M_{Z_2|Z_2^*}$, where $Z_{i2}^* = \frac{1}{\sum_{k \in A_i^*} |C_k^*|} \sum_{j \in C_k^*} X_j$. Z_{i2}^* presents the average characteristics of individual ks' direct links in the latent contextual network, where individual ks are direct links of individual i in the latent peer-effect network. Since Z_{i2}^* is also an average term of characteristics of individuals in network C, we can make use of the previous results. Similarly, for convenience, we suppress the subscript i.

For simplicity, we assume $|A_i| = 1$, that is in the latent peer-effect network i only has one direct link. Based on the comparison between the latent network and its two measurements, there are four cases similar to the discussion of Z_1. Notice, it is enough to consider two cases that $|A_i| = |A_i^*|$ and $|A_i| > |A_i^*|$.

Case 1: $|A_i| = |A_i^*| = 1$

Under this case, we observe one link in network A but we do not know if the link is the true link in the latent network. If we observe the right link, say individual l, then $Pr(Z_2 = \frac{k}{m}|Z_2^* = \frac{k^*}{n})$ is the same as $Pr(Z_{l1} = \frac{k}{m}|Z_{l1}^* = \frac{k^*}{n})$. However,if we observe the wrong link, say some individual l' other than l, then $Pr(Z_2 = \frac{k}{m}|Z_2^* = \frac{k^*}{n})$ is the same as $Pr(Z_{l'1} = \frac{k}{m}|Z_{l'1}^* = \frac{k^*}{n})$. Recall we assume the probability of observing the right link in network A is given by $Pr(A_{ij}=1|A^*_{ij}=1)=Pr(A_{ij}=0|A^*_{ij}=0)=\lambda$.

$$Pr(Z_2=\frac{k}{m}|Z^*_2=\frac{k^*}{n})=\lambda Pr(Z_{l1}=\frac{k}{m}|Z^*_{l1}=\frac{k^*}{n})+(1-\lambda)Pr(Z_{l'1}=\frac{k}{m}|Z^*_{l'1}=\frac{k^*}{n})$$

Case 2: $|A_i|=2>|A^*_i|=1$

Under this case, we observe two links in network A but there is only one true link in the latent network A^*. There are two possible outcomes. We may observe the true link, say individual l, and another wrong link, arbitrarily say l'. Or we may observe two wrong links, arbitrarily say l' and l''. Then $Pr(Z_2=\frac{k}{m}|Z^*_2=\frac{k^*}{n})$ follows the form:

$$\begin{aligned}Pr(Z_2=\frac{k}{m}|Z^*_2=\frac{k^*}{n})=&\,C^1_2\lambda(1-\lambda)Pr(Z_{l1}=\frac{k_l}{m_l}|Z^*_{l1}=\frac{k^*_l}{n_l})Pr(Z_{l'1}=\frac{k_{l'}}{m_{l'}}|Z^*_{l'1}=\frac{k^*_{l'}}{n_{l'}})\\&+(1-\lambda)^2Pr(Z_{l'1}=\frac{k_{l'}}{m_{l'}}|Z^*_{l'1}=\frac{k^*_{l'}}{n_{l'}})Pr(Z_{l''1}=\frac{k_{l''}}{m_{l''}}|Z^*_{l''1}=\frac{k^*_{l''}}{n_{l''}}),\end{aligned}$$

where

$$k_l+k_{l'}=k_{l'}+k_{l''}=k,$$

$$k^*_l+k^*_{l'}=k^*_{l'}+k^*_{l''}=k^*,$$

$$n_l+n_{l'}=n_{l'}+n_{l''}=n,$$

$$m_l+m_{l'}=m_{l'}+m_{l''}=m,$$

For general case, where $|A^*_i|=n_a$, $|A_i|=m_a$, and $m_a\geq n_a$, we can write $Pr(Z_2=\frac{k}{m}|Z^*_2=\frac{k^*}{n})$ in the following form:

$$Pr(Z_2 = \frac{k}{m} | Z_2^* = \frac{k^*}{n}) = \sum_{\omega=0}^{n_a} \lambda^{n_a-\omega}(1-\lambda)^{m_a-(n_a-\omega)} C_{m_a}^{n_a-\omega} \Bigg[\prod_{l=1}^{n_a-\omega} C_{n_a}^{n_a-\omega} Pr(Z_{l1} = \frac{k_l}{m_l} | Z_{l1}^* = \frac{k_l^*}{n_l}) \prod_{l'=n_a+1}^{n_a+1+\omega} Pr(Z_{l'1} = \frac{k_{l'}}{m_{l'}} | Z_{l'1}^* = \frac{k_{l'}^*}{n_{l'}}) \Bigg],$$

where for each ω such that $0 \leq \omega \leq n_a$,

$$\sum_{l=1}^{\omega} k_l + \sum_{l'=n_a+1}^{\omega+n_a+1} k_{l'} = k,$$

$$\sum_{l=1}^{\omega} k_l^* + \sum_{l'=n_a+1}^{\omega+n_a+1} k_{l'}^* = k^*,$$

$$\sum_{l=1}^{\omega} n_l + \sum_{l'=n_a+1}^{\omega+n_a+1} n_{l'} = n,$$

$$\sum_{l=1}^{\omega} m_l + \sum_{l'=n_a+1}^{\omega+n_a+1} m_{l'} = m.$$

Notice that in the above equation, we suppress subscript i. Without loss of generality, we can order all the individuals in the network such that the first n_a individuals are the true links to some arbitrary individual i. For example, individual $l \in \{1,2,...,n_a\}$ is a direct link to i in the latent network A^*.

APPENDIX B

Proof of Proposition 3.1

Proof. First, we assume that X^c is one-dimensional and consider two sets of structural parameters θ and θ', where $\theta = (\alpha, \beta, \gamma, \delta)$ and $\theta' = (\alpha', \beta', \gamma', \delta')$. Suppose θ and θ' lead to the same reduced form, which means

$$\alpha(I - \beta A^*)^{-1}\iota = \alpha'(I - \beta' A^*)^{-1}\iota, \tag{5.1}$$

$$\gamma(I - \beta A^*)^{-1} = \gamma'(I - \beta' A^*)^{-1}, \tag{5.2}$$

and

$$\delta(I - \beta A^*)^{-1}C^* = \delta'(I - \beta' A^*)^{-1}C^*. \tag{5.3}$$

We want to show that these imply $\theta = \theta'$. Pre-multiply (5.2) and (5.3) by $(I - \beta A^*)(I - \beta' A^*)$. Since for all b, $(I - bA^*)^{-1}A^* = A^*(I - bA^*)^{-1}$ (Both $(I - bA^*)^{-1}$ and A^* are symmetric matrix), we have

$$\gamma(I - \beta' A^*) = \gamma'(I - \beta A^*),$$

and

$$\delta(I - \beta' A^*)C^* = \delta'(I - \beta A^*)C^*.$$

To rearrange the above two equations, we get

$$(\gamma - \gamma')I + (\gamma'\beta - \gamma\beta')A^* = 0 \tag{5.4}$$

and

$$(\delta - \delta')C^* + (\delta'\beta - \delta\beta')A^*C^* = 0. \tag{5.5}$$

Since the matrices I, A^*, C^* and A^*C^* are linearly independent, (5.4) and (5.5) implies

$$\gamma = \gamma',\ \delta = \delta',\ \beta\gamma' = \beta'\gamma,\ \beta\delta' = \beta'\delta. \tag{5.6}$$

If $\delta \neq 0$, we know that $\beta = \beta'$ by the last equation in (5.6). Furthermore, observe that (5.1) implies $\alpha\iota = \alpha'\iota$, hence $\alpha = \alpha'$. Therefore, we have $\theta = \theta'$.

Next, suppose that there are K characteristics in X^c and that parameters γ_k is associated with characteristics k. Then, θ and θ' leading to the same reduced form implies (5.1), (5.3) and

$$\gamma_k(I - \beta A^*)^{-1} = \gamma'_k(I - \beta' A^*)^{-1}, \tag{5.7}$$

for all k. Similarly by previous arguments, for all k, we can pre-multiply (5.7) by $(I - \beta A^*)(I - \beta' A^*)$. Then we get $\alpha = \alpha'$, $\beta = \beta'$, $\gamma_k = \gamma'_k$ for all k and $\delta = \delta'$, and hence $\theta = \theta'$. □

Proof of Lemma 3.1

Proof. The proof is provided in Corollary 1 and 2 in Blume et al. (2015) and hence is omitted. □

Proof of Equation 3.7

Proof.

$$Equation(3.4) \Rightarrow m(X_i^c, Z_i^*; \theta) \equiv E(Y_i | X_i^c, Z_i^*)$$

$$
\begin{aligned}
\int m(X_i^c, Z_i^*; \theta) f_{Z^*|Z,X^c} dZ^* &= \int E(Y_i|X_i^c, Z_i^*) f_{Z^*|Z,X^c} dZ^* \\
&= \int \int Y_i f_{Y|Z^*,X^c} dY f_{Z^*|Z,X^c} dZ^* \\
&= \int Y_i \int f_{Y|Z^*,X^c} f_{Z^*|Z,X^c} dZ^* dY \\
&= \int Y_i \int f_{Y|Z^*,X^c} \frac{f_{Z^*,Z,X^c}}{f_{Z,X^c}} dZ^* dY \\
&= \int \frac{Y_i}{f_{Z,X^c}} \int f_{Y|Z^*,X^c} f_{Z^*,Z,X^c} dZ^* dY
\end{aligned}
$$

By Assumption 4.2.(ii),

$$
\begin{aligned}
\int \frac{Y_i}{f_{Z,X^c}} \int f_{Y|Z^*,X^c} f_{Z^*,Z,X^c} dZ^* dY &= \int \frac{Y_i}{f_{Z,X^c}} \int f_{Y|Z^*,Z,X^c} f_{Z^*,Z,X^c} dZ^* dY \\
&= \int Y_i \frac{f_{Y,Z,X^c}}{f_{Z,X^c}} dY \\
&= E(Y_i|X^c, Z)
\end{aligned}
$$

$\square$

Proof of Lemma 3.2

Proof. By Assumption 3.5, the rank of $D_{Z_m^*|X^c}$ is K_m. By the rank inequality, for any $p \times n$ matrix A and $n \times q$ matrix B,

$$
\text{rank}(A) + \text{rank}(B) - n \leq \text{rank}(AB) \leq \min\{\text{rank}(A), \text{rank}(B)\}.
$$

Then we can show that $M_{Z_m|Z_m^*,X^c} D_{Z_m^*|X^c}$ has rank K_m. By applying the inequality again, we can conclude that the matrix

$$
M_{Z_m,Z_m'|X^c} = M_{Z_m|Z_m^*,X^c} D_{Z_m^*|X^c} M_{Z_m'|Z_m^*,X^c}^T.
$$

has rank K_m. $\square$

Proof of Proposition 3.2

Proof. Assumptions 3.4-3.6 implies that the observed matrix $M_{Z_m,Z'_m,Y|X^c}M^{-1}_{Z_m,Z'_m|X^c}$ admits an eigenvalue-eigenvector decomposition that

$$M_{Z_m,Z'_m,Y|X^c}M^{-1}_{Z_m,Z'_m|X^c} = M_{Z_m|Z^*_m,X^c}D_{y|Z^*_m,X^c}M^{-1}_{Z_m|Z^*_m,X^c}.$$

Assumption 3.7 (i) ensures that there are no duplicate values for the diagonal elements in $D_{Y|Z^*_m,X^c}$, which correspond to $f_{Y|Z^*_m,X^c}$. Therefore, the eigenvectors are linearly independent with each other. The next step is to determine which eigenvectors corresponds to $f_{Z_m|Z^*_m,X^c}(\cdot|j,x^c)$ for $j = 1,2,...,K_m$. Condition (ii) imposes an ordering for the eigenvectors. Consequently, we can identify $f_{Z_m|Z^*_m,X^c}$ by checking the location of the largest value in each eigenvector.

The conditional distribution of Z_m can be represented as $f_{Z_m|X^c} = \sum_{Z^*_m} f_{Z_m|Z^*_m,X^c} f_{Z^*_m|X^c}$. The matrix representation is

$$M_{Z_m|X^c} = M_{Z_m|Z^*_m,X^c}M_{Z^*_m|X^c},$$

Therefore, the conditional density of Z^*_m can be identified as

$$M_{Z^*_m|X^c} = M^{-1}_{Z_m|Z^*_m,X^c}M_{Z_m|X^c},\ m = 1,2.$$

Note that the invertibility of $M_{Z_m|Z^*_m,X^c}$ is guaranteed by Assumption 3.6. $\square$

Proof of Theorem 3.1

Proof. The proof follows the discussion in the text and hence is omitted. $\square$

Proof of Proposition 3.3

Proof. First, we show that $\widehat{f}_{Z_m,Z'_m,X_c,Y} \xrightarrow{p} f_{Z_m,Z'_m,X_c,Y}$. Fix $n \in \mathbb{N}_+$, $z \in \mathcal{Z}$, $z' \in \mathcal{Z}'$, $x \in \mathcal{X}^c$ and $y \in \mathcal{Y}$. By Chebychev's inequality, there exists $c > 0$ such that

$$Pr\left(\left|\widehat{f}_{Z_m,Z'_m,X_c,Y}(z,z',x,y) - \mathbb{E}\left[\widehat{f}_{Z_m,Z'_m,X_c,Y}(z,z',x,y)\right]\right| \geq c\right) \leq c^2 \operatorname{Var}\left[\widehat{f}_{Z_m,Z'_m,X_c,Y}(z,z',x,y)\right].$$

Since $\widehat{f}_{Z_m,Z'_m,X_c,Y}(z,z',x,y) = \frac{1}{n}\sum_i \mathbb{1}(Z_{mi} = z, Z'_{mi} = z', X_i^c = x, Y_i = y) \equiv \frac{1}{n}\sum_i \mathbb{1}_i$,

$$\begin{aligned}\operatorname{Var}\left[\frac{1}{n}\sum_{i=1}^{n}\mathbb{1}(Z_{mi} = z, Z'_{mi} = z', X_i^c = x, Y_i = y)\right] &= \frac{1}{n^2}\sum_{i=1}^{n}\operatorname{Var}\mathbb{1}_i + \frac{1}{n^2}\sum_{i=1}^{n}\sum_{j\neq i}^{n}\operatorname{Cov}(\mathbb{1}_i, \mathbb{1}_j)\\ &= \frac{1}{n}\operatorname{Var}\mathbb{1}_i + \frac{n(n-1)}{n^2}\operatorname{Cov}(\mathbb{1}_i, \mathbb{1}_j)\\ &= o(1)\end{aligned}$$

by Assumption 3.9 (iii). Furthermore, by Assumption 3.9 (i)-(ii) we have

$$\mathbb{E}\left[\widehat{f}_{Z_m,Z'_m,X_c,Y}(z,z',x,y)\right] = \mathbb{E}\left(\frac{1}{n}\sum_{i=1}^{n}\mathbb{1}_i\right) = \mathbb{E}(\mathbb{1}_i) = f_{Z_m,Z'_m,X_c,Y}.$$

Therefore, $\widehat{f}_{Z_m,Z'_m,X_c,Y} \xrightarrow{p} f_{Z_m,Z'_m,X_c,Y}$. Similarly, we can show that $\widehat{f}_{Z_m,Z'_m,X_c} \xrightarrow{p} f_{Z_m,Z'_m,X_c}$ and $\widehat{f}_{Z_m,X_c} \xrightarrow{p} f_{Z_m,X_c}$. Furthermore, $\widehat{f}_{X^c} \xrightarrow{p} f_{X_c}$ by law of large numbers. Hence, by continuous mapping theorem $\widehat{f}_{Z_m,Z'_m,Y|X_c} \xrightarrow{p} f_{Z_m,Z'_m,Y|X_c}$, $\widehat{f}_{Z_m,Z'_m|X_c} \xrightarrow{p} f_{Z_m,Z'_m|X_c}$ and $\widehat{f}_{Z_m|X_c} \xrightarrow{p} f_{Z_m|X_c}$.

Since the eigenvector function $\psi(\cdot)$ is a well-behaved analytical function (Andrew, Chu and Lancaster 1993), we have

$$\widehat{f}_{Z_m|Z_m^*,X^c} \xrightarrow{p} f_{Z_m|Z_m^*,X^c}$$

for $m = 1,2$. Finally, by continuous mapping theorem we can show that

$$\widehat{f}_{Z_m^*|X^c} \xrightarrow{p} f_{Z_m^*|X^c},\ m = 1,2.$$

$\square$

Proof of Proposition 3.4

Proof. We prove this proposition by verifying conditions (i)-(iv) of Theorem 2.1 in Newey and McFadden (1994). It is easy to see that conditions (ii) and (iii) are satisfied by Assumption 3.10. Therefore, we only need verify conditions (i) and (iv), i.e., $\mathbb{E}[\log c_\sigma(F_1,F_2)]$ is uniquely maximized at σ_0 and

$$\sup_{\sigma\in\Sigma}\left|\frac{1}{n}\sum_{i=1}^{n}\log\left[c_\sigma(\widehat{F}_1(Z_{1i}|X_i^c),\widehat{F}_2(Z_{2i}|X_i^c))\right]-\mathbb{E}[\log c_\sigma(F_1(Z_{1i}|X_i^c),F_2(Z_{2i}|X_i^c))]\right|=o_p(1). \quad (5.8)$$

First, $\mathbb{E}[\log c_\sigma(F(Z_{1i}),F(Z_{2i}))]$ is uniquely maximized at σ_0 by Sklar's theorem and Lemma 2.2 of Newey and McFadden (1994). Then, in order to show (5.8), we follow the proof of Theorm 4.2 in Wooldridge (1994) and show that

$$\sup_{(\sigma,F_1,F_2)\in\Omega}\left|\frac{1}{n}\sum_{i=1}^{n}\log\left[c_\sigma(F_1(Z_{1i}|X_i^c),F_2(Z_{1i}|X_i^c))\right]-\mathbb{E}[\log c_\sigma(F_1(Z_{1i}|X_i^c),F_2(Z_{2i}|X_i^c))]\right|=o_p(1). \quad (5.9)$$

Let $\delta>0$ be a positive number to be set later. Since Ω is compact, there exists a finite covering of Ω, say $B_\delta(\sigma_j,F_{1j},F_{2j})$, $j=1,2,...,K(\delta)$, where $B_\delta(\sigma_j,F_{1j},F_{2j})$ is the sphere of radius δ about (σ_j,F_{1j},F_{2j}), i.e.,

$$B_\delta(\sigma_j,F_{1j},F_{2j})=\left\{(\sigma,F_1,F_2)\in\Omega|\ \|(\sigma,F_1,F_2)-(\sigma_j,F_{1j},F_{2j})\|<\delta\right\}.$$

Set $\omega=(\sigma,F_1,F_2)$, $B_j=B_\delta(\omega_j)$, $L_n(\omega)=1/n\sum_{i=1}^n\log\left[c_\sigma(F_1(Z_{1i}|X_i^c),F_2(Z_{2i}|X_i^c))\right]$, $\overline{L}_n(\omega)=\mathbb{E}[L_n(\omega)]$. Since $\Omega\subset\bigcup_{j=1}^K B_\delta(\omega_j)$, it follows that

$$\begin{aligned}Pr\left[\sup_{\omega\in\Omega}\left|L_n(\omega)-\overline{L}_n(\omega)\right|>\varepsilon\right]&\le Pr\left[\max_{1\le j\le K(\delta)}\sup_{\omega\in B_j}\left|L_n(\omega)-\overline{L}_n(\omega)\right|>\varepsilon\right]\\ &\le\sum_{j=1}^{K(\delta)}Pr\left[\sup_{\omega\in B_j}\left|L_n(\omega)-\overline{L}_n(\omega)\right|>\varepsilon\right]. \quad (5.10)\end{aligned}$$

We will show that each probability in the summand is $o(1)$.

Define $l_i(\omega) = \log c_\sigma(F_1(Z_{1i}|X_i^c), F_2(Z_{2i}|X_i^c))$, $\bar{l}_i(\omega) = \mathbb{E}[l_i(\omega)]$. For $\omega \in B_j$,

$$\begin{aligned}
&\left|L_n(\omega) - \overline{L}_n(\omega)\right| \leq \left|L_n(\omega) - L_n(\omega_j)\right| + \left|L_n(\omega_j) - \overline{L}_n(\omega_j)\right| + \left|\overline{L}_n(\omega_j) - \overline{L}_n(\omega)\right| \\
\leq &\frac{1}{n}\sum_{i=1}^{n}\left|l_i(\omega) - l_i(\omega_j)\right| + \left|\frac{1}{n}\sum_{i=1}^{n} l_i(\omega_j) - \bar{l}_i(\omega_j)\right| + \frac{1}{n}\sum_{i=1}^{n}\left|\bar{l}_i(\omega) - \bar{l}_i(\omega_j)\right|
\end{aligned}$$

by the triangle inequality. By Assumption 3.10 (iii) and the monotonicity of expectation, for $\omega \in B_j$,

$$\left|l_i(\omega) - l_i(\omega_j)\right| \leq c_i\|\omega - \omega_j\| \leq \delta c_i$$

and

$$\left|\bar{l}_i(\omega) - \bar{l}_i(\omega_j)\right| \leq \bar{c}_i\|\omega - \omega_j\| \leq \delta\bar{c}_i,$$

where $c_i = \sup_{\omega\in B_j}\|\nabla_\omega l_i(\omega)\|$ and $\bar{c}_i = \mathbb{E}(c_i)$. Therefore, we have

$$\begin{aligned}
\sup_{\omega\in B_j}\left|L_n(\omega) - \overline{L}_n(\omega)\right| &\leq \delta\left(\frac{1}{n}\sum_{i=1}^{n} c_i + \frac{1}{n}\sum_{i=1}^{n}\bar{c}_i\right) + \left|\frac{1}{n}\sum_{i=1}^{n} l_i(\omega_j) - \bar{l}_i(\omega_j)\right| \\
&\leq 2\delta\frac{1}{n}\sum_{i=1}^{n}\bar{c}_i + \delta\left|\frac{1}{n}\sum_{i=1}^{n} c_i - \frac{1}{n}\sum_{i=1}^{n}\bar{c}_i\right| + \left|\frac{1}{n}\sum_{i=1}^{n} l_i(\omega_j) - \bar{l}_i(\omega_j)\right| \\
&= 2\delta\bar{c}_i + \delta\left|\frac{1}{n}\sum_{i=1}^{n} c_i - \bar{c}_i\right| + \left|\frac{1}{n}\sum_{i=1}^{n} l_i(\omega_j) - \bar{l}_i(\omega_j)\right|
\end{aligned}$$

where the second inequality is by triangle inequality and the last equality is by Assumption 3.9. Since $\bar{c}_i \leq \overline{C} < \infty$ by Assumption 3.10 (iii). It follows that

$$Pr\left[\sup_{\omega\in B_j}\left|L_n(\omega) - \overline{L}_n(\omega)\right| > \varepsilon\right] \leq Pr\left[\delta\left|\frac{1}{n}\sum_{i=1}^{n} c_i - \bar{c}_i\right| + \left|\frac{1}{n}\sum_{i=1}^{n} l_i(\omega_j) - \bar{l}_i(\omega_j)\right| > \varepsilon - 2\delta\overline{C}\right].$$

Now choose $\delta \leq 1$ such that $\varepsilon - 2\delta\overline{C} > \varepsilon/2$. Then

$$Pr\left[\sup_{\omega\in B_j}\left|L_n(\omega)-\overline{L}_n(\omega)\right| > \varepsilon\right] \leq Pr\left[\left|\frac{1}{n}\sum_{i=1}^{n} c_i - \overline{c}_i\right| + \left|\frac{1}{n}\sum_{i=1}^{n} l_i(\omega_j) - \overline{l}_i(\omega_j)\right| > \frac{\varepsilon}{2}\right].$$

Next, choose N so that

$$Pr\left[\left|\frac{1}{n}\sum_{i=1}^{n} c_i - \overline{c}_i\right| + \left|\frac{1}{n}\sum_{i=1}^{n} l_i(\omega_j) - \overline{l}_i(\omega_j)\right| > \frac{\varepsilon}{2}\right] \leq \frac{\varepsilon}{K(\delta)}, \tag{5.11}$$

for all $n \geq N$ and all $j = 1, 2, ..., K(\delta)$. Note that this is possible because

$$\left|\frac{1}{n}\sum_{i=1}^{n} c_i - \overline{c}_i\right| = o_p(1)$$

and

$$\left|\frac{1}{n}\sum_{i=1}^{n} l_i(\omega_j) - \overline{l}_i(\omega_j)\right| = o_p(1)$$

by similar argument as in the proof of Proposition 3.3. Therefore, (5.10) and (5.11) together imply that for all $n \geq N$ and $\varepsilon > 0$

$$Pr\left[\sup_{\omega\in\Omega}\left|L_n(\omega)-\overline{L}_n(\omega)\right| > \varepsilon\right] \leq \varepsilon,$$

which proves (5.9). Finally, since $F_n(Z_{li}) \xrightarrow{p} F(Z_{li})$, $l = 1, 2$ by similar argument as in the proof of Proposition 3.3, (5.9) implies that

$$\sup_{\sigma\in\Sigma}\left|\frac{1}{n}\sum_{i=1}^{n}\log\left[c_\sigma(\widehat{F}_1(Z_{1i}|X_i^c), \widehat{F}_2(Z_{2i}|X_i^c))\right] - \mathbb{E}[\log c_\sigma(F_1(Z_{1i}|X_i^c), F_2(Z_{2i}|X_i^c))]\right| = o_p(1)$$

by Theorem 3.7 of White (1996). Consequently, $\widehat{\sigma} \xrightarrow{p} \sigma_0$ by Theorem 2.1 of Newey and McFadden (1994). $\square$

Proof of Proposition 3.5

Proof. By Assumption 3.11 (i) and Proposition 3.4, $\widehat{\sigma}$ is an interior point of Σ with probability approaching 1 as $n \to \infty$. Therefore, with probability approaching 1, $\nabla_\sigma L_n(\widehat{\sigma}, \widehat{F}_1, \widehat{F}_2) = 0$. By Taylor expansion at σ_0, we have

$$\nabla_\sigma L_n(\sigma_0, \widehat{F}_1, \widehat{F}_2) + \nabla_{\sigma,\sigma} L_n(\sigma^*, \widehat{F}_1, \widehat{F}_2)(\widehat{\sigma} - \sigma_0) = 0 \text{ w.p.a } 1,$$

where σ^* lies between $\widehat{\sigma}$ and σ_0. Consequently,

$$\sqrt{n}(\widehat{\sigma} - \sigma_0) = \frac{\sqrt{n}\nabla_\sigma L_n(\sigma_0, \widehat{F}_1, \widehat{F}_2)}{-\nabla_{\sigma,\sigma} L_n(\sigma^*, \widehat{F}_1, \widehat{F}_2)}.$$

Now, we prove this proposition in three steps:

Step 1. First, we will show that $\sqrt{n}\nabla_\sigma L_n(\sigma_0, \widehat{F}_1, \widehat{F}_2) \xrightarrow{d} N(0, v^2)$ by similar argument as in the proof of Theorem 8.1 in Newey and McFadden (1994). By Assumption 3.10 (iii) and Taylor expansion at $(F_{Z_1,X^c}, F_{Z_1,X^c}, X^c)$, we have

$$\begin{aligned}\sqrt{n}\nabla_\sigma L_n(\sigma_0, \widehat{F}_1, \widehat{F}_2) &= \sqrt{n}\nabla_\sigma L_n(\sigma_0, F_1, F_2) + \sqrt{n}\nabla_{\sigma,F_{Z_1,X^c}} L_n(\sigma_0, F_1, F_2)(\widehat{F}_{Z_1,X^c} - F_{Z_1,X^c}) \\ &+ \sqrt{n}\nabla_{\sigma,F_{Z_2,X^c}} L_n(\sigma_0, F_1, F_2)(\widehat{F}_{Z_2,X^c} - F_{Z_2,X^c}) + \sqrt{n}\nabla_{\sigma,F_{X^c}} L_n(\sigma_0, F_1, F_2)(\widehat{F}_{X^c} - F_{X^c}) + r,\end{aligned}$$

where $r = O_p\left(\sqrt{n}\|\widehat{F} - F\|_\infty^2\right)$. Therefore, by Assumption 3.11 (iii) and Donsker's theorem,

$$\begin{aligned}&|\sqrt{n}\nabla_\sigma L_n(\sigma_0, \widehat{F}_1, \widehat{F}_2) - \sqrt{n}\nabla_\sigma L_n(\sigma_0, F_1, F_2) - \sqrt{n}\nabla_{\sigma,F_{Z_1,X^c}} L_n(\sigma_0, F_1, F_2)(\widehat{F}_{Z_1,X^c} - F_{Z_1,X^c}) \\ &- \sqrt{n}\nabla_{\sigma,F_{Z_2,X^c}} L_n(\sigma_0, F_1, F_2)(\widehat{F}_{Z_2,X^c} - F_{Z_2,X^c}) - \sqrt{n}\nabla_{\sigma,F_{X^c}} L_n(\sigma_0, F_1, F_2)(\widehat{F}_{X^c} - F_{X^c})| = o_p(1).\end{aligned}$$

Hence, the remainder term from the linearization is small. Next, we verify the stochastic equicontinuity condition (Andrews 1994) for $\sqrt{n}\nabla_{\sigma,F_{Z_m,X^c}} L_n(\sigma_0, F_1, F_2)(\widehat{F}_{Z_m,X^c} - F_{Z_m,X^c})$ and $\sqrt{n}\nabla_{\sigma,F_{X^c}} L_n(\sigma_0, F_1, F_2)(\widehat{F}_{X^c} - F_{X^c})$. By Assumption 3.11 (iv) and Theorem 4.2 of Rio (2017), it is clear that

$$
\begin{aligned}
&\sqrt{n}\left[\nabla_{\sigma,F_{Z_m,X^c}}L_n(\sigma_0,F_1,F_2)-\mathbb{E}(\nabla_{\sigma,F_{Z_m,X^c}}L_n(\sigma_0,F_1,F_2))\right](\widehat{F}_{Z_m,X^c}-F_{Z_m,X^c})\\
=&\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\left\{\nabla_{\sigma,F_{Z_m,X^c}}l_i(\sigma_0,F_1,F_2)-\mathbb{E}[\nabla_{\sigma,F_{Z_m,X^c}}l_i(\sigma_0,F_1,F_2)]\right\}(\widehat{F}_{Z_m,X^c}-F_{Z_m,X^c})=o_p(1).
\end{aligned}
$$

Similarly, we can show that

$$
\sqrt{n}\left[\nabla_{\sigma,F_{X^c}}L_n(\sigma_0,F_1,F_2)-\mathbb{E}(\nabla_{\sigma,F_{X^c}}L_n(\sigma_0,F_1,F_2))\right](\widehat{F}_{X^c}-F_{X^c})=o_p(1).
$$

Define F_n to be the empirical measure. Then,

$$
\begin{aligned}
&\int\nabla_{\sigma,F_{Z_m,X^c}}l_i(\sigma_0,F_1,F_2)(\widehat{F}_{Z_m,X^c}-F_{Z_m,X^c})dF\\
=&\int\nabla_{\sigma,F_{Z_m,X^c}}l_i(\sigma_0,F_1,F_2))\left[\mathbb{1}(Z_{mi}\leq z_m,X_i^c\leq x^c)-F_{Z_m,X^c}(z,x^c)\right]dF_n,
\end{aligned}
$$

and

$$
\int\nabla_{\sigma,F_{X^c}}l_i(\sigma_0,F_1,F_2)(\widehat{F}_{X^c}-F_{X^c})dF=\int\nabla_{\sigma,F_{X^c}}l_i(\sigma_0,F_1,F_2))\left[\mathbb{1}(X_i^c\leq x^c)-F_{X^c}(x^c)\right]dF_n,
$$

Consequently, condition (iii) of Theorem 8.1 in Newey and McFadden (1994) is verified. Finally, their condition (iv) is trivially satisfied. Therefore,

$$
\sqrt{n}\nabla_\sigma L_n(\sigma_0,\widehat{F}_1,\widehat{F}_2)\xrightarrow{d}N(0,v^2)
$$

by CLT for weakly dependent random variable (see, e.g., Theorem 4.2 of Rio 2017), where

$$
v^2=\lim_{n\to\infty}\frac{1}{n}\mathrm{Var}\left\{\sum_{i=1}^{n}[\nabla_\sigma\log C_\delta(F_1(Z_{1i}|X_i^c),F(Z_{2i}|X^c))+K_1(Z_{1i},X_i^c)+K_2(Z_{2i},X_i^c)+K_3(X_i^c)]\right\}.
$$

Step 2. Next, we show that $\nabla_{\sigma,\sigma} L_n(\sigma_0, \widehat{F}_1, \widehat{F}_2) \xrightarrow{p} b$. We have

$$\nabla_{\sigma,\sigma} L_n(\sigma^*, \widehat{F}_1, \widehat{F}_2) = \frac{1}{n} \sum_{i=1}^{n} \nabla_{\sigma,\sigma} l_i(\sigma^*, \widehat{F}_1, \widehat{F}_2).$$

By argument similar to the proof of Proposition 3.4, $\nabla_{\sigma,\sigma} l_i(\cdot,\cdot,\cdot)$ satisfies the uniform weak law of large numbers. Furthermore, $\sigma^* \xrightarrow{p} \sigma_0$ and $\widehat{F}_m \xrightarrow{p} F_m$, $m = 1,2$. Hence, by Theorem 3.7 of White (1996),

$$\nabla_{\sigma,\sigma} L_n(\sigma^*, \widehat{F}_1, \widehat{F}_2) \xrightarrow{p} \mathbb{E}[\nabla_{\sigma,\sigma} \log C_\delta(F_1, F_2)] = b.$$

Step 3. Finally, by combining the results in Steps 1 and 2 and applying Slutsky's theorem, we can conclude that

$$\sqrt{n}(\widehat{\sigma} - \sigma_0) \xrightarrow{p} N(0, v^2/b^2)$$

□

Proof of Theorem 3.2

Proof. We prove this theorem by verifying conditions (i)-(iv) of Theorem 2.1 in Newey and McFadden (1994). It is easy to see that conditions (ii) and (iii) are satisfied by Assumption 3.12. Furthermore, condition (i) is satisfied by the identification result. Therefore, we only need to verify condition (iv), i.e.,

$$\sup_{\theta\in\Theta} \left| \frac{1}{n} \sum_{i=1}^{n} g_i(\theta, \widehat{\sigma}, \widehat{\eta}) - \mathbb{E}[g_i(\theta, \sigma, \eta)] \right|,$$

which can be proved similarly to Proposition 3.4. Therefore, we can conclude that

$$\widehat{\theta} \xrightarrow{p} \theta_0.$$

□

Proof of Theorem 3.3

Proof. By Assumption 3.13 (i) and Theorem 3.2, $\widehat{\theta}$ is an interior point of Θ with probability approaching 1 as $n \to \infty$. Therefore, with probability approaching 1 the first order condition is

$$\left[\frac{1}{n}\sum_{i=1}^{n}\nabla_{\theta} g_i(\widehat{\theta},\widehat{\sigma},\widehat{\eta})\right]^T \widehat{V}\left[\frac{1}{n}\sum_{i=1}^{n} g_i(\widehat{\theta},\widehat{\sigma},\widehat{\eta})\right] = 0.$$

By Taylor expansion at θ_0, we have

$$\frac{1}{n}\sum_{i=1}^{n} g_i(\widehat{\theta},\widehat{\sigma},\widehat{\eta}) = \frac{1}{n} g_i(\theta_0,\widehat{\sigma},\widehat{\eta}) + \frac{1}{n}\nabla_{\theta} g_i(\theta^*,\widehat{\sigma},\widehat{\eta})(\widehat{\theta}-\theta_0) \text{ w.p.a 1},$$

where θ^* is between $\widehat{\theta}$ and θ_0. Consequently,

$$\begin{aligned}\sqrt{n}(\widehat{\theta}-\theta_0) = &-\left\{\left[\frac{1}{n}\sum_{i=1}^{n}\nabla_{\theta} g_i(\widehat{\theta},\widehat{\sigma},\widehat{\eta})\right]^T \widehat{V}\left[\frac{1}{n}\sum_{i=1}^{n}\nabla_{\theta} g_i(\theta^*,\widehat{\sigma},\widehat{\eta})\right]\right\}^{-1} \\ &\times\left[\frac{1}{n}\sum_{i=1}^{n}\nabla_{\theta} g_i(\widehat{\theta},\widehat{\sigma},\widehat{\eta})\right]^T \widehat{V}\left[\frac{1}{\sqrt{n}}\sum_{i=1}^{n} g_i(\theta_0,\widehat{\sigma},\widehat{\eta})\right]\end{aligned}$$

Now, we prove this theorem in three steps:

Step 1. First, we will show that $1/\sqrt{n}\sum_{i=1}^{n} g_i(\theta_0,\widehat{\sigma},\widehat{\eta}) \xrightarrow{d} N(0,\Lambda)$ by similar argument as in the proof of Theorems 6.1 and 8.1 in Newey and McFadden (1994). By Taylor expansion at (σ_0,η_0),

$$\begin{aligned}\frac{1}{\sqrt{n}}\sum_{i=1}^{n} g_i(\theta_0,\widehat{\sigma},\widehat{\eta}) = &\frac{1}{\sqrt{n}}\sum_{i=1}^{n} g_i(\theta_0,\sigma_0,\eta_0) + \frac{1}{\sqrt{n}}\sum_{i=1}^{n}\nabla_{\sigma} g_i(\theta_0,\sigma_0,\eta_0)(\widehat{\sigma}-\sigma_0) \\ &+ \frac{1}{\sqrt{n}}\sum_{i=1}^{n}\nabla_{\eta} g_i(\theta_0,\sigma_0,\eta_0)(\widehat{\eta}-\eta_0) + r,\end{aligned}$$

where the remainder term r is $o_p(1)$ by Proposition 3.5 and Assumption 3.13 (iii). First, note that

$$\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\nabla_{\sigma} g_i(\theta_0,\sigma_0,\eta_0)(\widehat{\sigma}-\sigma_0) = \frac{1}{n}\sum_{i=1}^{n}\nabla_{\sigma} g_i(\theta_0,\sigma_0,\eta_0)\sqrt{n}(\widehat{\sigma}-\sigma_0)$$

By argument similar to the proof of Proposition 3.4, $\nabla_{\sigma} g_i(\cdot,\cdot,\cdot)$ satisfies the uniform weak law of large numbers. Besides, $\widehat{\sigma}$ is an asymptotically linear estimator with influence function

$$\frac{\nabla_\sigma \log C_\delta(F_1(Z_{1i}|X_i^c), F_2(Z_{2i}|X^c)) + K_1(Z_{1i}, X_i^c) + K_2(Z_{2i}, X_i^c) + K_3(X_i^c)}{-\mathbb{E}[\nabla_{\sigma,\sigma} \log C_\delta(F_1, F_2)]}$$

by Proposition 3.5. Therefore, by continuous mapping theorem

$$\begin{aligned}
&\frac{1}{\sqrt{n}}\sum_{i=1}^n \nabla_\sigma g_i(\theta_0,\sigma_0,\eta_0)(\widehat{\sigma}-\sigma_0)\\
=&\mathbb{E}[\nabla_\sigma g_i(\theta_0,\sigma_0,\eta_0)]\\
&\frac{1}{\sqrt{n}}\sum_{i=1}^n \frac{\nabla_\sigma \log C_\delta(F_1(Z_{1i}|X_i^c), F_2(Z_{2i}|X^c)) + K_1(Z_{1i}, X_i^c) + K_2(Z_{2i}, X_i^c) + K_3(X_i^c)}{-\mathbb{E}[\nabla_{\sigma,\sigma} \log C_\delta(F_1, F_2)]}\\
&+o_p(1)\\
=&\frac{1}{\sqrt{n}}\sum_{i=1}^n \nu_i + o_p(1).
\end{aligned}$$

Next, we show that

$$\frac{1}{\sqrt{n}}\sum_{i=1}^n \nabla_\eta g_i(\theta_0,\sigma_0,\eta_0)(\widehat{\eta}-\eta_0) = \frac{1}{\sqrt{n}}\sum_{i=1}^n \xi_i + o_p(1).$$

Following the proof of Theorem 8.1 in Newey and McFadden (1994), we verify the stochastic equicontinuity condition for $1/\sqrt{n}\sum_{i=1}^n \nabla_\eta g_i(\theta_0,\sigma_0,\eta_0)(\widehat{\eta}-\eta_0)$. By Assumption 3.11 (iv), Assumption 3.12 (v) and Theorem 4.2 of Rio (2017), it is clear that

$$1/\sqrt{n}\sum_{i=1}^n \nabla_\eta g_i(\theta_0,\sigma_0,\eta_0)(\widehat{\eta}-\eta_0) = o_p(1).$$

Furthermore, we have

$$\begin{aligned}
&\int \nabla_{f_{Z_m,Z'_m,Y,X^c}} g_i(\theta_0,\sigma_0,\eta_0)(\widehat{f}_{Z_m,Z'_m,Y,X^c} - f_{Z_m,Z'_m,Y,X^c})dF \\
=&\int \left[\mathbb{1}(\zeta_{mi}=\zeta_m) - f_{Z_m,Z'_m,Y,X^c}(\zeta_m)\right] \nabla_{f_{Z_m,Z'_m,Y,X^c}} g_i(\theta_0,\sigma_0,\eta_0)dF_n.
\end{aligned}$$

Similar results can be shown for

$$\int \nabla_{f_{Z_m,Z'_m,X^c}} g_i(\theta_0,\sigma_0,\eta_0)(\widehat{f}_{Z_m,Z'_m,X^c} - f_{Z_m,Z'_m,X^c})dF,$$

$$\int \nabla_{f_{Z_m,X^c}} g_i(\theta_0,\sigma_0,\eta_0)(\widehat{f}_{Z_m,X^c} - f_{Z_m,X^c})dF,$$

$$\int \nabla_{f_{Z,X^c}} g_i(\theta_0,\sigma_0,\eta_0)(\widehat{f}_{Z,X^c} - f_{Z,X^c})dF$$

and

$$\int \nabla_{f_{X^c}} g_i(\theta_0,\sigma_0,\eta_0)(\widehat{f}_{X^c} - f_{X^c})dF$$

for $m = 1,2$. Consequently, condition (iii) of Theorem 8.1 in Newey and McFadden (1994) is verified. Finally, their condition (iv) is trivially satisfied. Therefore,

$$\frac{1}{\sqrt{n}}\sum_{i=1}^{n} \nabla_\eta g_i(\theta_0,\sigma_0,\eta_0)(\widehat{\eta}-\eta_0) = \frac{1}{\sqrt{n}}\sum_{i=1}^{n}\xi_i + o_p(1),$$

and hence

$$\frac{1}{\sqrt{n}}\sum_{i=1}^{n} g_i(\theta_0,\widehat{\sigma},\widehat{\eta}) \xrightarrow{d} N(0,\Lambda)$$

by CLT for weakly dependent random variable (see, e.g., Theorem 4.2 of Rio 2017), where

$$\Lambda = \lim_{n\to\infty}\frac{1}{n}\mathrm{Var}\left\{\sum_{i=1}^{n}\left[g_i(\theta_0,\sigma_0,\eta_0)+\nu_i+\xi_i\right]\right\}.$$

Step 2. Next, we show that $1/n\sum_{i=1}^{n}\nabla_\theta g_i(\widehat{\theta},\widehat{\sigma},\widehat{\eta}) \xrightarrow{p} R$. By argument similar to the proof of

Proposition 3.4, $\nabla_\theta g_i(\cdot,\cdot,\cdot)$ satisfies the uniform weak law of large numbers. Moreover, $\widehat{\theta} \xrightarrow{p} \theta_0$, $\widehat{\sigma} \xrightarrow{p} \sigma_0$ and $\widehat{\eta} \xrightarrow{p} \eta_0$. Hence, by Theorem 3.7 of White (1996),

$$\frac{1}{n}\sum_{i=1}^{n} \nabla_\theta g_i(\widehat{\theta},\widehat{\sigma},\widehat{\eta}) \xrightarrow{p} \mathbb{E}[\nabla_\theta g_i(\theta_0,\sigma_0,\eta_0)] = R.$$

Similarly, we can show that

$$\frac{1}{n}\sum_{i=1}^{n} \nabla_\theta g_i(\theta^*,\widehat{\sigma},\widehat{\eta}) \xrightarrow{p} R.$$

Consequently,

$$\begin{aligned}
&\left\{\left[\frac{1}{n}\sum_{i=1}^{n} \nabla_\theta g_i(\widehat{\theta},\widehat{\sigma},\widehat{\eta})\right]^T \widehat{V} \left[\frac{1}{n}\sum_{i=1}^{n} \nabla_\theta g_i(\theta^*,\widehat{\sigma},\widehat{\eta})\right]\right\}^{-1} \\
&\xrightarrow{p} \left\{\mathbb{E}[\nabla_\theta g_i(\theta_0,\sigma_0,\eta_0)]^T V \mathbb{E}[\nabla_\theta g_i(\theta_0,\sigma_0,\eta_0)]\right\}^{-1} \\
&= G^{-1}
\end{aligned}$$

by continuous mapping theorem.

Step 3. Finally, by combining the results in Steps 1 and 2 and applying Slutsky's theorem, we can conclude that

$$\sqrt{n}(\widehat{\theta}-\theta_0) \xrightarrow{d} N(0, G^{-1}R^T V \Lambda V R G^{-1}).$$

□