# *Guanxi* and Income Inequality

ZAICHAO DU, JIANNAN GUO AND LIN ZHANG*

Southwestern University of Finance and Economics

May 1, 2013

## Abstract

*Guanxi*, loosely translated as "connections", is the core structure of Chinese society. This paper proposes a measure of *guanxi* to capture its multiple-dimensional nature and studies its impact on income inequality, using China household finance survey data. In line with the seminal theory of Fei Xiaotong (1992, From the Soil: the Foundations of Chinese Society), principal components analysis selects three main components to construct *Guanxi* Index: static relationships, authority and power, and means to maintain/develop connections. This paper also shows that *guanxi*, besides wealth concentration and human capital, is a key factor determining income inequality, and becomes quantitatively more important than human capital for the top 25 percent of urban households.

**Keywords**: *Guanxi* index; social networks; income inequality; China; principal components analysis.

**JEL Classifications**: D85; Z13; D30; C81.

* Zaichao Du, Jiannan Guo, and Lin Zhang are at Research Institute of Economics and Management of Southwestern University of Finance and Economics. Contact information: Zaichao Du, duzc@swufe.edu.cn; Jiannan Guo, guojiannan@hotmail.com; Lin Zhang, zhanglin@swufe.edu.cn. We have benefited from the research assistance of Jincai He. Zaichao Du gratefully acknowledges the research grant from Ministry of Education of P.R. China.

# 1 Introduction

Many observers and scholars recognize that *guanxi*, loosely translated as "connections", is the core structure of Chinese society, e.g. Liang (1949), Jacobs (1979, 1980), Fei (1992), and Gold et al. (2004), among others. As pointed out by the founder of social science in China Fei Xiaotong (1992), Chinese society is composed by overlapping ripples or *guanxi* that is "*formed from a stone thrown into a lake, each circle spreading out from the center becomes more distant and at the same time more insignificant*". In such a *guanxi* based society where the economic transformation is associated with rising income inequality to an unprecedented level: the top 10 percent of households' income share is 57 percent in 2010, the role of *guanxi* for the rising income inequality has caught special attention.

The purpose of this paper is to propose a measure of *guanxi* and study its contribution to the income inequality of Chinese households. We note that *guanxi* is a multidimensional concept, studied in the literature as special relationship, connection, social capital, or exchange, and existing measures can only capture one or two dimensions of *guanxi*. For instance, *guanxi* is measured as social capital in Knight and Yue (2008), and Zhang and Li (2003), and as kinship networks in Peng (2004). We also note that existing studies on Chinese income inequality are centered on examining factors to explain the rising urban-rural gap during the economic transformation. These factors include political strategies that favor certain industries and sectors (Yang, 1999; Yang and Zhou, 1999; Kanbur and Zhang, 2005), development of financial sectors (Zhang, 2004), dual economy structure between state economy and staple agriculture (Putterman, 1992), urban-rural human capital gap (Guo, 2005), opening-up policy (Wei and Zhao, 2012), and education and occupation from an aspect of individual characteristics (Su et al. 2013). *Guanxi*, the core structure of Chinese society, seems to be overlooked by the existing studies on income inequality. Yet, in the market economy, as declared by He (1998), *guanxi* has an inordinate role as a major mean not only for the redistribution of resources, but also

the accumulation of wealth and the rechanneling of public property into private hands.

In this paper, we construct a measure of *guanxi* capturing its multiple dimensions, using China Household Finance Survey (CHFS) data. Principal components analysis selects three main components to construct *Guanxi* Index: static relationships, authority and power, and means to maintain/develop connections. This measure is in line with the seminal theory of Fei (1992) that each household is at the center, its authority and power determines the size of its *guanxi*, and expanding ripples out from the center relies on means to maintain/develop connections.

We also show that besides physical capital and human capital, *guanxi* is another key contributor to income inequality, and becomes quantitatively more important than human capital for the top 25 percent of urban households. These results are robust to different data analysis methods and estimation specifications. As far as we know, we are the first in the literature to propose a measure of *guanxi* that captures its multiple dimensions, and to rigorously quantify the contribution of *guanxi* to income inequality in China.

Formally, we proceed as follows. First, we examine household finance survey data that randomly selects 8438 households from 2585 counties in China. Based on both theories and empirical studies on *guanxi*, we extract 23 *guanxi* related variables after carefully considering the issue of missing values. Through cluster analysis, principal components analysis, and factor analysis, 9 variables are selected to construct *guanxi* index. These variables are characterized as 3 items measuring static relationships, authority and power, and means to maintain/develop connections. In particular, the variable on the number of brothers and sisters of the husband and wife of each household is selected for static relationships, reflecting the blood base /social base of each family. Variables such as communist party membership, job position and job type are chosen to measure the social status or the authority and power of households. Variables such as expenditure on gift to other families or individuals, business dinning expenditure, communication expenditure, local transportation expenditure and entertainment expenditure are selected to measure means to maintain/develop

connections. We conduct extensive reliability test of the index and find that the index is reliable.

Second, to estimate the impact of *guanxi* on income inequality, we carefully construct measures of wealth and human capital. The family wealth includes the financial assets, fixed assets, as well as the value of land and agricultural machinery for farmers. Human capital includes average education of workers (and education squared), health status dummies of the household head and spouse, and the average professional skills of workers. In recognition of possible endogeneity between income and *guanxi*, we use income data of 2011 and other explanatory variables of 2010. Our results show that for 1 percent increase of the *Guanxi* index the household income will increase by 5.86 percent, which is about the same magnitude as the return for education. The return of *guanxi* in rural China is much bigger than that in urban.

Third, considering the income equation may be different across different quantiles, we then run quantile regressions at 10, 25, 50, 75 and 90 percent quantiles, respectively. We obtain that for the whole sample and urban, the returns of *Guanxi* for high quantiles are higher than low quantiles in general. For rural households, however, the returns of *Guanxi* for lower 25 percent and median are higher than higher quantiles, which in turn are higher than the lowest 10 percent.

Finally, to study the contributions of different variables to the income inequality in China, we use the Shapley value approach developed by Shorrocks (1999). Human capital accounts for about 30 percent of the income inequality in China. *Guanxi* index ranks as the third largest contributor to income inequalities, overtaking assets. It explains 19.10 percent of income inequality in urban China, and 12.02 percent in rural China. We also decompose the income inequality at different quantiles. We obtain that the higher the quantile is, the greater the contribution of *Guanxi* Index for the whole sample and the urban sample will be. *Guanxi* even overtakes human and assets as the second largest contributor to income inequality at 75 percent and 90 percent quantiles in urban China. *guanxi* is not the main contributor to inequality with each income quantile in rural China.

The paper is organized in the following way. In Section 2, we review the relevant literature. In Section 3, we construct *Guanxi* Index. In Section 4, we estimate the contribution of *guanxi* to income inequality. In Section 5, we conclude.

## 2 Review of the Existing Literature

We conduct the review of the existing literature in two ways. First, we review how existing studies conceptualize *guanxi*. Then, we review how previous empirical works measure *guanxi* and study its effects.

In the early studies, *guanxi* is defined as a special relationship between two persons, e.g. Jacobs (1979, 1980), Tsang (1994), and Yang (1994). Not only can it be applied to blood based relations such as family, kinships and in-law, it can also include social based relations such as relationship by nature (e.g. locality, classmate or alumni, teach-student, co-worker, neighbor, in the same profession), and relationship acquired (e.g. acquaintance, knowing the same person, friend, sworn brotherhood). However, defining *guanxi* as a special relationship captures only *guanxi* bases or its static dimension. *Guanxi* bases are neither necessary nor sufficient for producing *guanxi*. First, two strangers without any *guanxi* bases can still establish *guanxi*. Second, the existence of *guanxi* bases does not automatically lead to active *guanxi*. For instance, A and B were classmates twenty years ago and had been no contact after the graduation. They had *guanxi* base, but have no *guanxi*. Thus, maintain and develop *guanxi* or its dynamic dimension guaranty the existence of *guanxi*.

Some studies, defining *guanxi* as a network of social connections, capture its dynamic dimension, e.g. Liu (1983), and Gold, Guthrie and Wank (2004). Liu (1983) uses the electric circuit as a metaphor of *guanxi*, which can be connected and switched on. Getting involved with one person often means being involved with a whole network. Complex personal relationships with layer upon layer of interlocking connections form a dense net. The metaphor of electric circuit shares some similarities with the metaphor of ripples proposed by Fei (1949, 1992). Nevertheless,

Fei (1949, 1992) states that the center of the ripples is “self”, and that the size of ripples or *guanxi* depends on the center’s power and authority, and its efforts to expand the ripples.

With the development of social capital theory, another branch of studies define *guanxi* as a form of social investment or social capital that consists of networks of relationships and resources inherent in these networks, e.g. Butterfield (1983), Lin (2001), Xin and Pearce (1996) , Coleman (1988), and Tsang (1998). From the social capital dimension of *guanxi*, the analogy to develop and maintain *guanxi* is to put one’s money into a saving account or purchasing insurance policy so that one could get return or help whenever he/she needs. Each transaction will add or subtract the stock of social capital or the balance of favor. The dimension of *guanxi* as social capital can be naturally extended as organizational resources that not only affect firm performance across industries and regions but also create certain competitive advantages, e.g. Xin and Pearce (1996), Luo (1997), Tsang (1998), and Gu et al. (2008), among others.

However, *guanxi* is the core structure of Chinese society. As stated by Fei (1949, 1992), Chinese society cannot be adequately conceptualized in terms of organizations, nor the West in terms of *guanxi*. Fei (1949, 1992) uses a metaphor of haystacks to conceptualize Western society. The rice straw (i.e. individual) is bound into small bundles. Several bundles are bound into larger bundles, which in turn make up a stack. The separate straws, the separate bundles, and finally the separate stacks all fit together to make up the whole haystack (i.e. organization). In this way, the separately bound bundles can be stacked in an orderly way. Western social structure consists of functional organizations or groups with clear boundaries. Membership in these groups is unambiguous so that everyone knows who is and who is not a member. The rights and duties of members are clearly delineated. In effect, Chinese society cannot be adequately conceptualized in terms of haystacks, nor the West in terms of ripples in water.

According to Fei (1949, 1992), the origins of two modes of organization rely on

primitive tribal formations and cultures. In primitive nomadic economies, living together and cooperate were a precondition of life. But in a settled agricultural society, everyone earns his or her own living from the land and feels the need of companions only under fortuitous, temporary, or special circumstances. To these people, starting relationships with others is a matter of secondary importance. In addition, he traces the Western organizational mode to Christianity, and the differential mode of Chinese society to Confucian ethics1.

As *guanxi* is the core structure of Chinese society, the art of *guanxi* or *guanxi* practice is of central importance to Chinese people. Some studies define *guanxi* practice as the reciprocal exchange between two persons for a specific purpose, e.g. Hwang (1987), Yang (1989, 2002), and Yeung and Tung (1996). Guthrie (1998) delineates the difference between *guanxi* and *guanxi* practice: *guanxi* are social connections, while *guanxi* practice is the use of these social connections for specific ends. In other words, *guanxi* practice or gift economy studies how to develop and maintain *guanxi* for specific ends. In line with Fei (1949, 1992), he emphasizes that expand ripples or *guanxi* needs to tui (i.e. push) it from the center, or needs to develop and maintain it.

With multiple dimensions, it is a challenging task to measure *guanxi*. In the early studies *guanxi* was introduced to business audience as a cultural phenomenon and empirical studies were centered on its impact on firm performance, e.g. Abramson and Ai (1999), Luo (1997), Tsang (1998), and Gu et al. (2008). In these studies, *guanxi* is directly measured with responses, obtained from in-depth interviews with managers, on the importance of *guanxi* and the use of *guanxi*. The same approach is used by more systematic studies carried by researchers in the areas of anthropology, sociology, and psychology, e.g. Jacobs (1979), Hwang (1987), and Yang (1989, 2002).

Some empirical works, carried by researchers in economics and sociology, study the impacts of *guanxi* on entrepreneurship, employment, and income, e.g. Knight and Yue (2008), Zhang and Li (2003), Bian (1994), and Peng (2004). These studies

[1] See more details in Fei (1937, 1992, Chapter 4 )

consider *guanxi* as social capital or kinship networks, and use indicators from survey data to measure *guanxi*. Knight and Yue (2008) study the impacts of social capital on labor income using a survey data of urban households from 6 provinces in 2000. *Guanxi* is measured with indicators such as communist party membership, parents communist membership, and social network (i.e. number of contacts, change gifts, and maintain contacts). They find that social capital contributes more to earned income than education. Peng (2004) finds that kinship networks have had large positive effects on the growth of private entrepreneurship using data of 366 Chinese villages from 22 counties. Kinship networks are measured by the proportion of households that belong to the largest lineage group (i.e. share the same surname) in the whole village. Zhang and Li (2003) find that *guanxi* or social networks had a positive effect on nonfarm employment using survey data covering 787 households from 6 counties in 1995. Social networks are measured by expenditure on gifts and qualitative indicators such as whether a worker: (i) has received help from family members or friends during the process of looking for a job; (ii) has received remittance from family members living in towns or cities; (iii) has family members working outside own region; (iv) has a family member as a local official.

Focusing on the effects of *guanxi* on income inequality, our study is related to the literature analyzing income inequality in China. With limited extent, we provide a selective review of literature closely related to our study.

It is widely accepted that urban-rural income gap and income inequality within rural/urban areas went up during the transition period (Khan and Riskin, 1998; Knight and Song, 2003; Benjamin et al. 2005; Heshmati, 2007, among others), and urban-rural gap explains a great part of the overall inequality (Yang, 1999; Wu and Perloff, 2005). As a result, existing studies are centered on examining factors to explain the rising urban-rural gap during the economic transformation. These factors include political strategies that favor certain industries and sectors (Yang, 1999; Yang and Zhou, 1999; Kanbur and Zhang, 2005), development of financial sectors (Zhang, 2004), dual economy structure between state economy and staple agriculture

(Putterman, 1992), urban-rural human capital gap (Guo, 2005), opening-up policy (Wei and Zhao, 2012).

Our study is closely related to Su and Heshmati (2013) that examine the dominants of income inequality from the aspect of individual characteristics using household survey data from nine provinces. They find that education and occupation are essential determinants of households' income inequality. Our study proposes a measure of *guanxi* and shows that *guanxi*, the core structure of Chinese society, is another key factor to explain income inequality.

In sum, *guanxi* is a multidimensional concept, studied in the literature as special relationship, connection, social capital, and exchange, and all these dimensions can be captured by the seminal theory of Fei (1949, 1992) on *guanxi*. Existing empirical studies can only measure one or two dimensions of *guanxi*. In our study, we propose a measure of *guanxi* to capture its multiple dimensions and study its impact on income inequality in China, using representative household survey data.

## 3 Construction of *Guanxi* Index

In this section, we first draw guidelines from both theoretic and empirical studies in selecting variables to measure *guanxi*. Second, we explain data sources and the definitions of *guanxi* related variables. Last, we present the construction of *Guanxi* index.

### 3.1 Select *guanxi* related variables

Because *guanxi* is multidimensional concept and it is impossible to use one variable to measure its multiple dimensions, constructing a composite index comes out as the only solution. Composite indicators have been widely recognised as a useful tool providing simple comparison to illustrate complex. The number of composite indicators cited is growing to more than 160 (Bandur 2006). They have unique advantage to simplify the analytical work to one index rather than identifying

common trends across many separate variables which are proven useful in benchmarking (Saltelli, 2007).

We use China Household Finance Survey data, conducted by Southwestern University of Finance and Economics in 2011. It randomly selects 80 counties among the total 2,585 counties in the country (Tibet, Xinjiang, Inner Mongolia, Macau and Hong Kong are not included). In each county, 4 communities are randomly selected. In total, there are 320 communities, from which 8438 households are randomly selected. It contains detailed information on households' demographic characteristics, assets and debts, insurance and social welfare, and income and expenditures.

From the above review of the literature on *guanxi*, we draw several guidelines in selecting variables into the pool before using index construction methods. First, to measure its static dimension as a special relation, variables measuring *guanxi* bases (i.e. blood and social bases) are included into the pool. We have two variables to measure *guanxi* bases: lineage group and number of siblings. Lineage group is defined as whether the principal or/and the spouse of the principal of a rural family belongs to the largest lineage group (i.e. share the same surname). Peng (2004) uses this variable to measure *guanxi*. The number of siblings of the principal and the spouse of the principal of a family measures the blood base of *guanxi*.

Second, variables measuring the power and authority of a household are crucial because they determine the size of *guanxi*. We include three variables: communist party membership, job position, and work unit type. Communist party membership is used by Knight and Yue (2008) and Morduch and Sicular (2000) to measure social capital. Job position is constructed by grade multiplied by the square root of the number of years on the position. Grade is assigned according to the hierarchy of job positions: 1 for group leader/ village cadre; 2 for (vice) chief of section / town cadre; 3 for (vice) director of a department; 4 for (vice) director of a bureau or higher positions. Work unit type refers to the type of the family members' work unit (i.e. government institutions or enterprises). Grade 6 is for government or military; 5 for publicly listed state-owned enterprises; 4 for public institutions; 3 for non-listed

state-owned enterprises; 2 for publicly listed non state-owned enterprises; 1 for non-listed private enterprises; 0 for other types of work units.

There are two different views in the literature regarding to whether *guanxi* and its role are declining or not in Chinese society as the economic transformation progresses. Guthrie (1998) claims, based on in-depth interviews[2] with Chinese officials and managers in Shanghai, that *guanxi* and the art of *guanxi* occupy a diminishing role in China's urban industrial economy as the economic transition progresses. On the other hand, Yang (2002) refutes that with the development of market economy, *guanxi* practice has moved out the area of the acquisition of consumption goods and provision of everyday needs during the planned economy period, and into the area of the acquisition of resources in the business world. According to Xin and Pearce (1996), in both planned and market economy, *guanxi* serves as mechanisms that reduce uncertainties (i.e. structure support) when public institutions and channels (i.e. structure protection) prove more or less dysfunctional.

Thus, in the planned economy, *guanxi* is a way of reversing the governmentalization of everyday life by redistributing consumption goods controlled by the state. In the market economy, with path dependence of North (1990), the government still has control over natural resources, access to bank loans, state contracts, favorable tax incentives, valuable market information, exemptions from laws and regulations, and so on. As a result, work unit type is ranked based the degree of government involvement. Higher involvement means greater power.

Third, to capture the dynamic dimension of *guanxi*, variables measuring how to develop and maintain *guanxi* should be in the pool. Several variables are included into the pool such as gift expenditure on non-family members, gift income from non-family members, expenditure on dining out, expenditure on entertainment, expenditure on communication, etc. Hwang (1987) observes that the common tactics to develop and maintain *guanxi* are being introduced by a third person, presenting a gift, and holding a feast or/and several entertainment activities for the other party.

[2] Together, Gutherie conducted 155 in-depth interviews, 81 of which were conducted on-site.

There are different principles of interaction and social treatment between individuals connected by different *guanxi* bases. According to Tsui and Farh (1997), the general rule of interaction between individuals having close kinship ties is unconditional protection which is rendered largely without anticipation of reciprocity. For individuals connected by long-distant relative base and social base, the general principle of interaction is that interpersonal favors and generosity are rendered with the anticipation of reciprocity.

Anticipation of reciprocity motivates the Chinese to do renqing, a norm of reciprocity and loosely translated as human debt to acquaintances, for another. Hwang (1987) argues that besides three justice norms used for social exchange (i.e. the equity, equality, and need rules) renqing, a variant of the universal equality rule, is prominent in Chinese society. The expressive tie (for example family) is governed by the need rule. The instrumental tie (for example temporary and unstable business relationship) is governed by the equity rule. The mixed tie (or a particularistic tie occurring among individuals connected by long-distant relative base and social base) is governed by renqing. When deciding to do someone a renqing or a favor, the allocator of renqing usually considers the cost of providing such renqing, the anticipation of reciprocation (positively correlated with the receiver's social position, resources owned, generous reputation in repaying favors), and *guanxi*.

Last, because social connections, perceived by others, are important factors in judging overall social status and social power, variables measuring how households manage the image of their social status should be in the pool too. We include several variables such as expenditure on luxury goods, number of cars owned by the family, car value, etc. Variables are defined in Table 1.

Table 1 Variable definition

| Variable | Definition |
|---|---|
| Lineage group | Whether the principal or/and the spouse of the principal of a rural family belongs to the largest lineage group (i.e. share the same surname) |
| Number of siblings | The number of siblings of the principal and the spouse of the principal of a |

| | |
|---|---|
| | family |
| Party membership | Whether P/S of a family is/are communist party member(s) |
| Job position | The family members' job positions. It is constructed by grade multiplied by the square root of the number of years on the position. Grade is assigned according to the hierarchy of job positions: 1 for group leader/ village cadre; 2 for (vice) chief of section / town cadre; 3 for (vice) director of a department; 4 for (vice) director of a bureau or higher positions. |
| Work unit type | The type of the family members' work unit (i.e. government institutions or enterprises). Grade 6 is for government or military; 5 for publicly listed state-owned enterprises; 4 for public institutions; 3 for non-listed state-owned enterprises; 2 for publicly listed non state-owned enterprises; 1 for non-listed private enterprises; 0 for other types of work units. |
| Gift exp. | Gift expenditure to non-family members |
| Gift income | Gift income from non-family members |
| Dining-out exp. | Annual family expenditure on dinning out. |
| Entertainment exp. | Annual family expenditure on KTV, bar, theatre, etc. |
| Communication exp. | Annual family expenditure on communication |
| Transportation exp. | Annual family expenditure on local transportation |
| Visiting exp. | Annual family expenditure on traveling to visit relatives and friends. |
| Play income | Income from playing mahjong and cards. |
| Luxury exp. | Annual family expenditure on luxury goods, calligraphy and paintings, etc. |
| Information source | Whether the family obtains information from relatives and friends. |
| Stock owned inf. | Whether family members work or used to work in the company that the family owns stocks |
| Stock decision | Whether relatives and friends help make stock purchasing or selling decisions. |
| Debt refused reasons | Whether debt application was refused because of no guarantor or no connections |
| Bank choice | Whether family choice of banks is based on *guanxi* |
| Channel debt | How a family seeks financing. 1 for borrowing from siblings; 2 from relatives; 3 from friends/colleagues. |
| Physical appearance | Physical appearance of the interviewee, ranked from 1 to 10, and 1 stands for bad looking 10 for good looking. |
| Number of cars owned | Number of cars owned by the family |
| Car grade | Evaluated car values of the community where the family resides. |

The descriptive statistics of the 23 *Guanxi* related variables is as follows:

Table 2 Statistics of *guanxi* related variables

| Variable | Obs | Mean | Std.Dev. | Min | Max |
|---|---|---|---|---|---|

| | | | | | |
|---|---|---|---|---|---|
| Gift exp. | 6151 | 8173 | 16918 | 0 | 472000 |
| Gift income | 4147 | 4298 | 13511 | 0 | 330000 |
| Lineage group | 4532 | 0.999 | 0.705 | 0 | 3 |
| Party membership | 8438 | 0.236 | 0.493 | 0 | 2 |
| Job position | 8438 | 0.532 | 2.064 | 0 | 22.14 |
| Work unit type | 8438 | 1.345 | 2.414 | 0 | 12 |
| Number of siblings | 8438 | 5.699 | 3.439 | 0 | 24 |
| Dining-out exp. | 8361 | 3616 | 24787 | 0 | 1.320e+06 |
| Transportation exp. | 8373 | 2052 | 5389 | 0 | 120000 |
| Communication exp. | 8386 | 1724 | 2217 | 0 | 48000 |
| Entertainment exp. | 8406 | 476.6 | 4728 | 0 | 288000 |
| play income | 123 | 6211 | 21507 | 0 | 200000 |
| Luxury exp. | 178 | 10074 | 4240 | 1.94 | 8.1818 |
| Visit exp. | 3674 | 3550 | 7818 | 3.64 | 181818 |
| Information source | 8438 | 0.324 | 0.468 | 0 | 1 |
| Stock owned inf. | 745 | 0.0322 | 0.177 | 0 | 1 |
| Stock decision | 745 | 0.0497 | 0.338 | 0 | 3 |
| Debt refused reasons | 169 | 0.266 | 0.668 | 0 | 2 |
| Bank choice | 348 | 0.147 | 0.491 | 0 | 2 |
| Channel debt | 2252 | 0.836 | 1.007 | 0 | 3 |
| Physical appearance | 8438 | 6.429 | 1.633 | 1 | 10 |
| Num of cars owned | 1225 | 1.11102 | .3844413 | 1 | 4 |
| Car grade | 3790 | 2.432 | 0.958 | 1 | 4 |

It is clear that some variables (Play income, Luxury exp., Stock owned information, Stock decision, Debt refused reasons, Bank choice, Channel debt, Number of cars owned) have too many (more than 6000) missing observations, which will largely affect the accuracy of the analysis. In fact, a trial-and-error process to find best candidates entering the composite index drops these 8 variables in the very first round. Considering the contribution to accumulative variation, 9 candidates are finally selected: Gift exp., Party membership, Job position, Work unit type, Number of siblings, Dining-out exp., Communication exp., and Transportation exp.

### 3.2 Data cleaning and suitability examinations

From Table 2 one can see that after dropping eight variables with more than 6000 missing observations, nearly all the variables have more than 8300 observations (out

of 8438 in total) except the variable 'Gift exp.', which has only 6240 valid responses – more than 26 percent are with missing value3.

In general there are three methods dealing with missing data: (1) case deletion, (2) single imputation or (3) multiple imputations. Case deletion simply omits the missing records from the analysis. This most straightforward method, however, may result in serious waste of information as it ignores possible systematic differences between complete and incomplete samples and may produce biased estimates if the cases are not missing at random. Moreover, standard errors will normally be larger in a reduced sample, given that less information is used. Little & Rubin (2002) summarise a rule of thumb that case deletion shall not be applied when a variable has more than 5 percent missing values.

In order to examine whether 'Gift exp.' is missing at random, a comparison of missing group (observations with missing value in 'Gift exp.') and non-missing group is made.

Table 3 Comparison of missing and non-missing groups

| | Missing group | | | Mean | Non-missing group | | |
|---|---|---|---|---|---|---|---|
| Variable | Obs | Mean | Std.Dev | comparison | Obs | Mean | Std.Dev |
| HH income | 2198 | 28162 | 87493 | < | 6240 | 47667 | 114966 |
| HH avg age | 2198 | 43.48 | 16.33 | > | 6240 | 40.29 | 14.42 |
| Gift income | 2198 | 1044 | 7395 | < | 6240 | 2525 | 10437 |
| Party membership | 2198 | 0.143 | 0.396 | < | 6240 | 0.269 | 0.519 |
| Job position | 2198 | 0.309 | 1.679 | < | 6240 | 0.611 | 2.178 |
| Work unit type | 2198 | 0.883 | 1.918 | < | 6240 | 1.507 | 2.547 |
| Dining-out exp. | 2198 | 2344 | 26686 | < | 6240 | 3570 | 20683 |
| Transportation exp. | 2198 | 1057 | 2865 | < | 6240 | 2132 | 5394 |
| Communication exp. | 2198 | 1073 | 1723 | < | 6240 | 1727 | 2078 |
| Risk attitude | 2196 | 4.212 | 1.413 | > | 6237 | 3.832 | 1.364 |
| Happiness | 2197 | 2.454 | 0.991 | > | 6239 | 2.252 | 0.820 |
| Num of rural obs | 1020 | 46.4 percent | | | 2224 | 35.65 | |
| Num of urban obs | 1178 | 53.6 percent | | | 4016 | 64.35 | |

On average, households in missing group are with lower income level and more

[3] In fact these records are not really 'missing'. A careful examination of the questionnaire shows that a question (G2001) is designed to ask 'whether the family has given more than 100 Yuan in value to any non-family member as gift in last one year', to which 2130 responds answered 'no'. Another 68 answered 'yes' but gave zero summation in sub-questions reporting detailed spending.

risk aversion. They report lower gift income as well. All other *Guanxi* variables show the same trend. It is then suspected that the two group are statistically different in different *guanxi* aspects, hence the value may not be missing at random. A Probit model is applied to test this hypothesis.

Table 4 Probit regression

| Dependant variable: whether 'Gift exp.' is reported as 'yes' | |
|---|---|
| Whether reporting 'Gift income' | 0.687*** |
| | (0.0324) |
| HH avg age | -0.00560*** |
| | (0.00111) |
| Party membership | 0.247*** |
| | (0.0384) |
| Number of siblings | 0.0284*** |
| | (0.00461) |
| Work unit type | 0.0352*** |
| | (0.00833) |
| Transportation exp. | 1.13e-05** |
| | (5.26e-06) |
| Communication exp. | 7.01e-05*** |
| | (1.10e-05) |
| Whether involved in business activities | -0.240*** |
| | (0.0520) |
| Whether holding stock accounts | -0.138** |
| | (0.0674) |
| Whether having cars | -0.0861 |
| | (0.0586) |
| Whether having fix-term deposit | -0.396*** |
| | (0.0458) |
| Intercept | 1.842*** |
| | (0.209) |
| N | 8,293 |

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

It is clear that all the *guanxi* variables used to explain the oddity are highly significant. Therefore some intrinsic relations must exist which in turn reject the hypothesis that the missing value come at random. Combining with the fact that more than 26 percent of cases are with missing value, imputing the missing cases becomes a must step.

Both single and multiple imputation methods are applied. In single imputation, the cases are divided into 15 groups with similar *guanxi* statistical characteristics. And mean value of each group is used to fill in the missing value in that group. Hence, the missing value of cases in the one group is imputed with the same mean value. In multiple imputation, Maximum Likelihood Estimation (MLE) is used to estimate the missing value one by one. Full analysis based on imputed data using both methods is conducted which shows similar pattern. Taking into account that single imputation may risk underestimating the variable when using same mean value to impute missing value in same group, MLE imputed data will be used in the construction of *Guanxi* index, while single imputed data will be kept to test the reliability of the results.

Four other variables – Dining-out exp., Transportation exp., Communication exp. and Entertainment exp. also have missing value in some records, although the number of these records is tiny comparing with the total sample size. Hence similar single imputation is conducted in order to keep most of the information from the sample.

The dataset after imputation contain 8438 observation of which all will be used in *Guanxi* index construction and later inequality analysis. These variables are normalised (standardised) to nil mean value and unity variance. KMO and SMC coefficients are calculated to verify whether these nine variables are suitable for principle component / factor analysis.

Table 5 Suitability analysis

| Variable | KMO | SMC |
|---|---|---|
| Gift exp. | 0.784 | 0.799 |
| Party membership | 0.689 | 0.581 |
| Job position | 0.664 | 0.669 |
| Work unit type | 0.702 | 0.457 |
| Number of siblings | 0.653 | 0.891 |
| Dining-out exp. | 0.612 | 0.871 |
| Communication exp. | 0.662 | 0.416 |
| Transport exp. | 0.679 | 0.479 |
| Entertainment exp. | 0.592 | 0.908 |
| Overall | 0.760 | |

Except for Entertainment exp. and Dining-out exp., all the other KMOs are above

0.65, while the overall KMO is 0.76, indicating the variables are acceptable for principle component / factor analysis. The high value of SMC leads to same conclusion. For the two variables with KMO around 0.6, however, the SMC are 0.908 and 0.871 respectively, suggesting the suitability to enter to enter a composite index.

The Cronbach Coefficient Alpha (henceforth c-alpha) (Cronbach, 1951) is also calculated as a supplementary examination – to provide a coefficient of reliability based on estimating the internal consistency of items – on how suitable these nine variable to enter PCA/FA analysis.

Table 6 C-Alpha of the nine variables

| Deleted individual indicator | Obs | Correlation with total | C-alpha |
|---|---|---|---|
| Gift exp. | 8438 | 0.432 | 0.691 |
| Party membership | 8438 | 0.462 | 0.682 |
| Job position | 8438 | 0.510 | 0.667 |
| Work unit type | 8438 | 0.527 | 0.661 |
| Number of siblings | 8438 | -0.338 | 0.717 |
| Dining-out exp. | 8438 | 0.501 | 0.670 |
| Communication exp. | 8438 | 0.614 | 0.632 |
| Transportation exp. | 8438 | 0.536 | 0.659 |
| Entertainment exp. | 8438 | 0.482 | 0.676 |
| Test scale | | | 0.702 |

Nunnally (1978) suggests 0.7 as an acceptable reliability threshold. Yet some authors are as lenient as to go to 0.6. In the table the overall scale is 0.702, meeting Nunnally (1978) standard. Communication exp. has the highest variable-total correlation and if deleted the coefficient alpha would be as low as 0.632. If Number of siblings were to be deleted from the set, the value of the standardized coefficient alpha would increase from the 0.66 to 0.72. Note that the same individual indicator has the lowest variable-total correlation value (-.338). This suggests that Number of siblings may represent different intrinsic of *guanxi* rather than other variables. Both examinations suggest the reliability of the nine variables for entering PCA/FA analysis.

### 3.3 Using PCA/FA to construct *Guanxi* Index

There are different ways and rigid processes constructing composite indicators. This

research combines the methods of principle component analysis and factor analysis to construct *Guanxi* index, and then refines the methodology in OECD (2008) for reliable tests on the index.

With standardisation of variables and default rotating rules, PCA and FA will both give same results. Hence PCA and FA will refer to same analysis in later discussion. Apply FA on cleaned dataset with standardised variables, the basic results are as follows:

Table 7 PCA/FA results

| Factors | *Eigenvalue* | *Proportion* | *Cumulative* |
|---|---|---|---|
| Factor1 | 2.256 | 0.251 | 0.251 |
| Factor2 | 1.505 | 0.167 | 0.418 |
| Factor3 | 1.099 | 0.122 | 0.540 |
| Factor4 | 1.000 | 0.111 | 0.651 |
| Factor5 | 0.828 | 0.0920 | 0.743 |
| Factor6 | 0.684 | 0.0760 | 0.819 |
| Factor7 | 0.585 | 0.0651 | 0.884 |
| Factor8 | 0.552 | 0.0613 | 0.946 |
| Factor9 | 0.490 | 0.0545 | 1 |

Factors with eigenvalue no less than unity are chosen as principle factors, which together explain 65 percent of total information. It is obvious that adding more factors as principle factor will contribute as less than 9 percent of explanation capability while suffering from losing more degree of freedom. The factor loading matrix is as follows.

Table 8 Factor loading matrix

| Variable | Factor 1 | Factor2 | Factor 3 | Factor4 |
|---|---|---|---|---|
| Gift exp. | 0.446 | -0.0362 | -0.466 | 0.188 |
| Party membership | 0.438 | 0.541 | 0.186 | 0.221 |
| Job position | 0.510 | 0.518 | 0.144 | 0.0397 |
| Work unit type | 0.519 | 0.561 | 0.213 | -0.120 |
| Number of siblings | -0.213 | 0.0445 | -0.0946 | 0.930 |
| Dining-out exp. | 0.499 | -0.508 | 0.454 | 0.119 |
| Communication exp | 0.687 | -0.197 | -0.356 | -0.0248 |
| Transportation exp. | 0.586 | -0.211 | -0.475 | -0.0712 |
| Entertainment exp. | 0.475 | -0.534 | 0.462 | 0.124 |

The factor loading explains how each variable contribute to a factor. In order to

have a clearer picture how such contribution works in economic sense – so as the factors can be reasonably named as a index – factor rotation is applied.

Table 9 Factors after rotation

| Factor | Variance | Proportion | Cumulative |
|---|---|---|---|
| Factor1 | 1.707 | 0.1896 | 0.1896 |
| Factor2 | 1.615 | 0.1794 | 0.3691 |
| Factor3 | 1.508 | 0.1675 | 0.5366 |
| Factor4 | 1.031 | 0.1145 | 0.6512 |

The rotation automatically keeps the first four factors which explains the same information as non-rotated ones.

Table 10 Factor Loading Matrix after rotation

| Variable | Factor 1 | Factor 2 | Factor 3 | Factor 4 |
|---|---|---|---|---|
| Gift exp. | 0.1033 | 0.6443 | -0.0206 | 0.1646 |
| Party membership | 0.7353 | 0.0463 | 0.0321 | 0.1544 |
| Job position | 0.7303 | 0.1233 | 0.0235 | -0.0304 |
| Work unit type | 0.7754 | 0.0649 | 0.0174 | -0.1941 |
| Number of siblings | -0.0357 | -0.0506 | -0.0307 | 0.9573 |
| Dining-out exp. | 0.0371 | 0.1086 | 0.8446 | -0.0292 |
| Communication exp. | 0.1314 | 0.7493 | 0.2236 | -0.0993 |
| Transportation exp. | 0.0237 | 0.7707 | 0.1035 | -0.1174 |
| Entertainment exp. | 0.0064 | 0.0931 | 0.8548 | -0.0231 |

Comparing with non-rotated loading matrix, the rotated one makes the dominant variables more obvious hence clearer conclusion can be drawn from it. In order to see relative contribution of variables in factor loading, the matrix is squared and scaled to unity.

Table 11 Squared factor loading (scaled to unity sum)

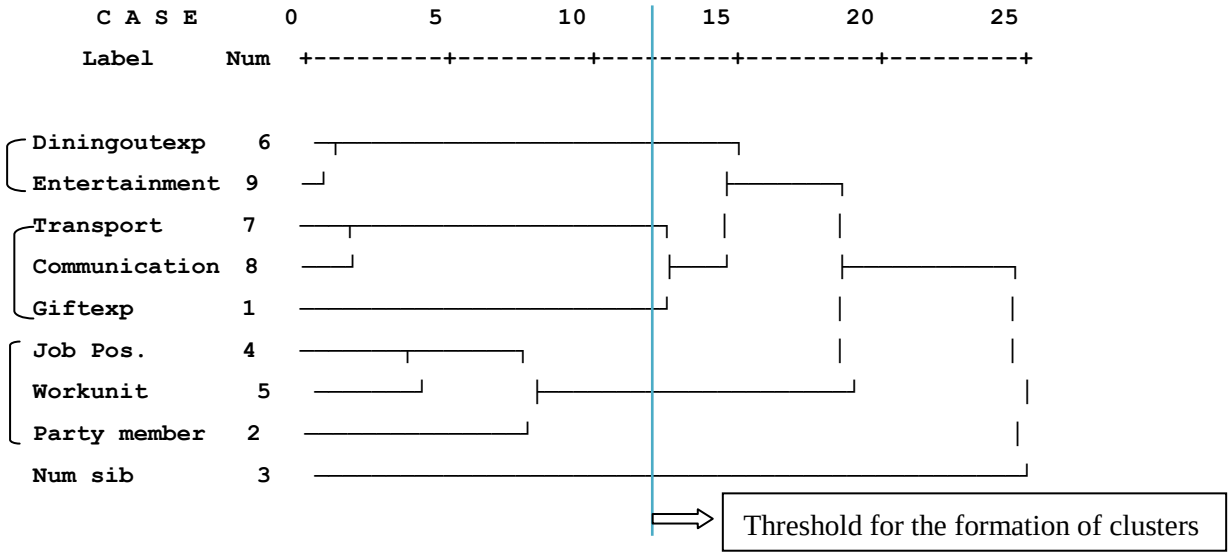
| Variable | Factor 1 | Factor 2 | Factor 3 | Factor 4 |
|---|---|---|---|---|
| Gift exp. | 0.01 | 0.26 | 0.00 | 0.03 |
| Party membership | 0.32 | 0.00 | 0.00 | 0.02 |
| Job position | 0.31 | 0.01 | 0.00 | 0.00 |
| Work unit type | 0.35 | 0.00 | 0.00 | 0.04 |
| Number of siblings | 0.00 | 0.00 | 0.00 | 0.89 |
| Dining-out exp. | 0.00 | 0.00 | 0.48 | 0.00 |
| Communication exp. | 0.01 | 0.35 | 0.03 | 0.01 |
| Transport exp. | 0.00 | 0.38 | 0.01 | 0.01 |
| Entertainment exp. | 0.00 | 0.00 | 0.48 | 0.00 |
| Sum of Column | 1 | 1 | 1 | 1 |

Now one stands better in explaining each factor.

In Factor 1 the variables contributing most (98 percent in total) are Party membership, Job position and Work unit type – three variables measuring social status that comes into being during one's career development. This status also stands for the social network resources as a composition of *guanxi* capital. Those dominating (99 percent in total) Factor 2 are Gift exp., Communication exp. and Transport exp., three variables that reflecting the dynamic feature of *guanxi* – it need to maintain as well as to invest. Factor 3 mainly dominated by Dining-out exp. and Entertainment exp. (98 percent contribution). This is another dynamic aspect of maintenance and investment of *guanxi*. Factor 4 has a single dominant contributor –Number of siblings. This factor shows the blood base of *guanxi*.

The explanation of these four factors can be verified by clustering analysis with the nine variables.

Table 12 Rescaled distance cluster combine




The clustering results show that:

- Party membership, Job position and Work unit type are one group, which are the same as dominant variables in Factor 1.
- Gift exp., Communication exp. and Transport exp. are one group, which are the same as dominant variables in Factor 2.
- Dining-out exp. and Entertainment exp. are one group, which are the same as dominant variables in Factor 3.

- Number of siblings, stands out to be one group, which is exactly the same as the only dominant contributor of Factor 4.

The four factors can be used as four sub-index that reflect different dimensions of *guanxi*. Combining the four a comprehensive *Guanxi* index can be finally produced, while the weights come from the relative contribution to the information explained:

| Variable | Factor 1 | Factor 2 | Factor 3 | Factor 4 |
|---|---|---|---|---|
| Explained Variance | 1.707 | 1.615 | 1.508 | 1.031 |
| Relative weight | 0.29 | 0.27 | 0.26 | 0.18 |

Therefore, the *Guanxi* index is given by :

Index=0.29*Factor1+0.27*Factor2+0.26*Factor3+0.18*Factor4

The descriptive statistics of *Guanxi* Index is

| Variable | Obs | Mean | Std.Dev. | Min | Max |
|---|---|---|---|---|---|
| Index | 8438 | 5.77e-11 | 0.507 | -0.679 | 14.36 |

If affine transformation is applied to map the index to the region of [0, 100], a 'standardised' *Guanxi* Index, index_100, can be produced:

| Variable | Obs | Mean | Std.Dev. | Min | Max |
|---|---|---|---|---|---|
| Index_100 | 8438 | 4.518 | 3.371 | 0 | 100 |

Further examination with Index, one may find that most of the observations (96 percent) fall in the region of [-1, 1].

| Region of Index | -1<=Index<=1 | 1<Index<=2 | 2<=Index<=3 | 3<=Index<=4 | 4<=Index<=5 |
|---|---|---|---|---|---|
| Observations | 8122 | 270 | 33 | 5 | 1 |

Hence, one can screen out the outliers to have a more condensed sample with 8122 observations. *Guanxi* Index is also calculated based on this condensed sample.

| Variable | Obs | Mean | Std.Dev. | Min | Max |
|---|---|---|---|---|---|
| index | 8122 | -0.063 | 0.331 | -0.679 | 0.993 |

Again, if the Index is projected to the region of [0, 100], standardised Index_100 can be produced:

| Variable | Obs | Mean | Std.Dev. | Min | Max |
|---|---|---|---|---|---|
| Index_100 | 8122 | 36.86 | 19.78 | 0 | 100 |

In sum, through cluster analysis, principal components analysis, and factor analysis, 9 variables are selected to construct *guanxi* index. These variables are characterized as 3 items measuring static relationships, authority and power, and means to maintain/develop connections. In particular, the variable on the number of siblings is selected for static relationships, reflecting the blood base of each family. Variables such as communist party membership, job position and work unit type are chosen to measure the social status or the authority and power of households. Variables such as gift expenditure, dinning-out expenditure, communication expenditure, local transportation expenditure and entertainment expenditure are selected to measure means to maintain/develop connections.

# 4 *Guanxi* and Income Inequality

In this section, we first estimate the income determination equation, and then conduct the income inequality decomposition.

## 4.1 Income determination equation

We focus on household income inequality in this study as in Lerman and Yitzhaki (1985), Morduch and Sicular (2002) and Wan and Zhou (2005) among many others. We start with the following household income determination equation

$$\ln(income_{ij})=\beta_0+\beta_1 Guanxi_{ij}+\beta_2 FC_{ij}+\beta_3 HC_{ij}+\beta_4 Assets_{ij}+c_i+u_{ij}, \quad (1)$$

where incomeij is the 2011 income of household j in community i. *Guanxi* is the constructed index; FC is the family characteristics, including household size, household size square, gender and the risk tolerance of the household head and the average age of workers (and age squared). The family human capital variables HC include average education of workers (and education squared), health status dummies of the household head and spouse and the average professional titles of workers. The

family Assets variables include the financial assets, fixed assets, as well as the value of land and agricultural machinery for farmers. Meanwhile, we include a community fixed effect term ci. To avoid the endogeneity problem, our dependent variable is the household income of 2011, while all the independent variables are observations in year 2010.

Table 13. Income determination equation

| | Whole Sample | Urban | Rural |
|---|---|---|---|
| VARIABLES | Ln_income | Ln_income | Ln_income |
| *Guanxi*_index | 0.0586*** | 0.0520*** | 0.0986*** |
| | (0.0111) | (0.0118) | (0.0282) |
| Hsize | 0.251*** | 0.184* | 0.253** |
| | (0.0796) | (0.108) | (0.113) |
| Square_ hsize | -0.0205*** | -0.0138 | -0.0216** |
| | (0.00784) | (0.0134) | (0.0104) |
| Gender | 0.0568* | 0.0710* | 0.0471* |
| | (0.0328) | (0.0380) | (0.0265) |
| Health_dummy1 | 0.190*** | 0.147** | 0.247*** |
| | (0.0527) | (0.0692) | (0.0821) |
| Health_dummy2 | 0.140** | 0.160** | 0.123* |
| | (0.0556) | (0.0749) | (0.0696) |
| Age | 0.0557*** | 0.0931*** | 0.0228 |
| | (0.0213) | (0.0293) | (0.0364) |
| Square_age | -0.000935*** | -0.00140*** | -0.000548 |
| | (0.000232) | (0.000341) | (0.000372) |
| Prof_ title | 0.186*** | 0.196*** | |
| | (0.0595) | (0.0621) | |
| Edu | 0.0669** | 0.0366* | 0.112** |
| | (0.0315) | (0.0212) | (0.0435) |
| Square_Edu | -0.00186 | -0.000342 | -0.00453 |
| | (0.00198) | (0.00230) | (0.00516) |
| Land_dummy | -0.223 | | -0.464 |
| | (0.162) | | (0.312) |
| Lvalue_land | 0.0237* | | 0.0412** |
| | (0.0142) | | (0.0204) |
| Lvalue_mach | 0.0269* | | 0.0197* |
| | (0.0160) | | (0.0116) |
| Risk_dummy | 0.0896* | 0.0181* | 0.199** |
| | (0.0494) | (0.0097) | (0.0829) |
| LNFasset_value | 0.00791** | 0.00955** | 0.00549 |
| | (0.00328) | (0.00376) | (0.00638) |
| LFasset_value | 0.0305*** | 0.0346*** | 0.0251* |

| | | | |
|---|---|---|---|
| | (0.00749) | (0.00907) | (0.0130) |
| Constant | 7.232*** | 7.038*** | 7.437*** |
| | (0.497) | (0.624) | (0.941) |
| Observations | 6,044 | 3,621 | 2,423 |
| R-squared | 0.429 | 0.326 | 0.286 |
| Number of ID1 | 320 | 161 | 159 |

Note: robust standard errors are reported in the parentheses. ***, ** and * represent significance at the level of 1 percent, 5 percent and 10 percent respectively, same for tables below.

Table 13 reports the estimation results for the income determination equation. Using all the sample data, we get that for 1 percent increase of the *Guanxi* index the household income will increase by 5.86 percent, which is about the same magnitude as the return for education. The return of *Guanxi* in rural China is much bigger than that in urban.

Considering the income equation may be different across different quantiles, we then run quantile regressions at 10, 25, 50, 75 and 90 percent quantiles, respectively. Table 14, 15 and 16 show the results for the whole sample, urban and rural, respectively. We see that for the whole sample and urban, the returns of *Guanxi* for high quantiles are higher than low quantiles in general. For rural households, however, the returns of *Guanxi* for lower 25 percent and median are higher than higher quantiles, which in turn are higher than the lowest 10 percent.

Table 14. Income equation quantile regression, whole sample

| VARIABLES | q10 | q25 | q50 | q75 | q90 |
|---|---|---|---|---|---|
| *Guanxi*_index | 0.0324*** | 0.0386*** | 0.0445*** | 0.0404*** | 0.0490*** |
| | (0.0111) | (0.00674) | (0.00577) | (0.00606) | (0.00702) |
| Hsize | 0.249* | 0.220*** | 0.127*** | 0.103*** | 0.0926* |
| | (0.128) | (0.0606) | (0.0329) | (0.0372) | (0.0482) |
| Square_ hsize | -0.0236 | -0.0154** | -0.00586* | -0.00469 | -0.00304 |
| | (0.0148) | (0.00650) | (0.00312) | (0.00367) | (0.00496) |
| Gender | 0.107* | 0.0946*** | 0.0521** | 0.0543* | 0.100*** |
| | (0.0579) | (0.0315) | (0.0256) | (0.0287) | (0.0341) |
| Health_dummy1 | 0.0848* | 0.0582** | 0.107*** | 0.112*** | 0.148*** |
| | (0.0466) | (0.0266) | (0.0224) | (0.0243) | (0.0277) |
| Health_dummy2 | 0.150*** | 0.0840*** | 0.0478** | 0.0235 | 0.000872 |
| | (0.0490) | (0.0287) | (0.0214) | (0.0222) | (0.0271) |
| Age | 0.200*** | 0.0438** | 0.0111* | -0.0143 | -0.0423*** |
| | (0.0691) | (0.0196) | (0.0059) | (0.0132) | (0.0140) |

| | | | | | |
|---|---|---|---|---|---|
| Square_Age | -0.00287*** | -0.000711*** | -0.000305** | 4.29e-05 | 0.000391** |
| | (0.000903) | (0.000234) | (0.000141) | (0.000154) | (0.000157) |
| Prof_ title | 0.207*** | 0.152*** | 0.0981*** | 0.0614** | 0.0473 |
| | (0.0372) | (0.0242) | (0.0213) | (0.0257) | (0.0303) |
| Edu | 0.0396*** | 0.0309** | 0.0249* | 0.00961* | 0.0138 |
| | (0.0126) | (0.0136) | (0.0144) | (0.0052) | (0.0176) |
| Square_Edu | -0.000878 | -0.000168 | 9.94e-05 | 0.000960 | 0.00260** |
| | (0.00220) | (0.00111) | (0.000787) | (0.000925) | (0.00113) |
| Land_dummy | -0.504*** | -0.0678 | 0.0277 | 0.0101 | 0.0599 |
| | (0.165) | (0.0918) | (0.0710) | (0.0755) | (0.0965) |
| Lvalue_land | 0.0352** | 0.000782 | 0.00271 | 0.000782 | 0.00560 |
| | (0.0154) | (0.00839) | (0.00647) | (0.00665) | (0.00903) |
| Lvalue_mach | 0.0476*** | 0.0195* | 0.00661 | 0.00459 | 0.00592 |
| | (0.0154) | (0.0106) | (0.00809) | (0.00812) | (0.00807) |
| Risk_dummy | 0.0496 | 0.0421* | 0.0613*** | 0.0669*** | 0.0836*** |
| | (0.0391) | (0.0220) | (0.0210) | (0.0235) | (0.0299) |
| Lvalue_NFasset | 0.00683*** | 0.00905*** | 0.00920*** | 0.0118*** | 0.0148*** |
| | (0.00251) | (0.00133) | (0.00138) | (0.00156) | (0.00194) |
| LFasset_value | 0.0412*** | 0.0258*** | 0.0193*** | 0.0167*** | 0.0136*** |
| | (0.00655) | (0.00395) | (0.00323) | (0.00346) | (0.00462) |
| Constant | 5.558 | 8.876*** | 10.00*** | 10.98*** | 11.74*** |
| | (3.564) | (0.441) | (0.319) | (0.280) | (0.333) |
| Observations | 6,044 | 6,044 | 6,044 | 6,044 | 6,044 |

Table 15. Income equation quantile regression, Urban

| VARIABLES | q10 | q25 | q50 | q75 | q90 |
|---|---|---|---|---|---|
| *Guanxi*_index | 0.0349*** | 0.0360*** | 0.0384*** | 0.0431*** | 0.0479*** |
| | (0.0101) | (0.00697) | (0.00679) | (0.00715) | (0.00729) |
| Hsize | 0.177* | 0.185*** | 0.111** | 0.0191 | 0.0769 |
| | (0.107) | (0.0620) | (0.0554) | (0.0599) | (0.0735) |
| Square_ hsize | -0.0152 | -0.0137** | -0.00402 | 0.00526 | -0.00215 |
| | (0.0124) | (0.00695) | (0.00632) | (0.00706) | (0.00810) |
| Gender | 0.114* | 0.111*** | 0.0960*** | 0.102*** | 0.147*** |
| | (0.0587) | (0.0328) | (0.0306) | (0.0319) | (0.0396) |
| Health_dummy1 | 0.0566 | 0.0195 | 0.0660*** | 0.0866*** | 0.156*** |
| | (0.0515) | (0.0334) | (0.0244) | (0.0260) | (0.0387) |
| Health_dummy2 | 0.161*** | 0.111*** | 0.0322 | 0.0153 | -0.0226 |
| | (0.0511) | (0.0337) | (0.0286) | (0.0332) | (0.0418) |
| Age | 0.132*** | 0.0462** | 0.00756 | -0.00335 | -0.0228 |
| | (0.0510) | (0.0180) | (0.0148) | (0.0146) | (0.0178) |
| Square_Age | -0.00194*** | -0.000742*** | -0.000268 | -0.000109 | 0.000138 |
| | (0.000679) | (0.000223) | (0.000182) | (0.000173) | (0.000204) |
| Prof_ title | 0.209*** | 0.159*** | 0.128*** | 0.0668*** | 0.0741** |
| | (0.0334) | (0.0248) | (0.0220) | (0.0258) | (0.0330) |

| | | | | | |
|---|---|---|---|---|---|
| Edu | 0.0289** | 0.00212* | 0.0181* | 0.00244* | 0.0239** |
| | (0.0129) | (0.0012) | (0.0097) | (0.00132) | (0.0107) |
| Square_Edu | -0.000642 | 0.00140 | 0.000184 | 0.00131 | 0.00270* |
| | (0.00236) | (0.00126) | (0.000729) | (0.00105) | (0.00149) |
| Risk_dummy | 0.0162 | 0.0252 | 0.0747*** | 0.0748*** | 0.112*** |
| | (0.0451) | (0.0276) | (0.0221) | (0.0261) | (0.0347) |
| Lvalue_NFasset | 0.00507** | 0.0101*** | 0.00942*** | 0.0107*** | 0.0143*** |
| | (0.00245) | (0.00156) | (0.00137) | (0.00183) | (0.00223) |
| LFasset_value | 0.0456*** | 0.0244*** | 0.0192*** | 0.0152*** | 0.0125*** |
| | (0.00805) | (0.00502) | (0.00368) | (0.00381) | (0.00448) |
| Constant | 6.744** | 9.073*** | 10.17*** | 10.99*** | 11.49*** |
| | (2.758) | (0.396) | (0.326) | (0.303) | (0.407) |
| Observations | 3,621 | 3,621 | 3,621 | 3,621 | 3,621 |

Table 16. Income equation quantile regression, Rural

| VARIABLES | q10 | q25 | q50 | q75 | q90 |
|---|---|---|---|---|---|
| *Guanxi*_index | 0.0382*** | 0.0610*** | 0.0545*** | 0.0473*** | 0.0486** |
| | (0.0122) | (0.0189) | (0.0138) | (0.0157) | (0.0189) |
| Hsize | 0.511 | 0.309* | 0.127* | 0.120* | 0.158** |
| | (0.428) | (0.172) | (0.0647) | (0.0690) | (0.0761) |
| Square_ hsize | -0.0502 | -0.0233 | -0.00554 | -0.00599 | -0.0101 |
| | (0.0478) | (0.0186) | (0.00826) | (0.00668) | (0.00697) |
| Gender | 0.0602* | 0.0788** | 0.0347* | 0.0372* | 0.0266*** |
| | (0.0348) | (0.0371) | (0.0213) | (0.0218) | (0.0085) |
| Health_dummy1 | 0.188* | 0.167** | 0.142*** | 0.150*** | 0.130** |
| | (0.105) | (0.0681) | (0.0407) | (0.0430) | (0.0523) |
| Health_dummy2 | 0.119** | 0.0172** | 0.0700* | 0.0239 | 0.0338 |
| | (0.0592) | (0.012) | (0.0365) | (0.0392) | (0.0450) |
| Age | 0.316** | 0.0362 | 0.0113 | -0.0154 | -0.0516* |
| | (0.148) | (0.0403) | (0.0168) | (0.0236) | (0.0267) |
| Square_Age | -0.00425** | -0.000588 | -0.000282 | 0.000104 | 0.000539* |
| | (0.00184) | (0.000474) | (0.000182) | (0.000256) | (0.000282) |
| Edu | 0.120*** | 0.0582** | 0.0533** | 0.0465*** | 0.00802 |
| | (0.0368) | (0.0261) | (0.0185) | (0.0143) | (0.0389) |
| Square_Edu | -0.00401 | -0.000557 | -0.00129 | -0.00172 | 0.00155 |
| | (0.00631) | (0.00374) | (0.00242) | (0.00296) | (0.00311) |
| Land_dummy | -1.084*** | -0.266 | -0.119 | -0.0702 | 0.115 |
| | (0.390) | (0.227) | (0.158) | (0.140) | (0.196) |
| Lvalue_land | 0.0868*** | 0.0140** | 0.0117 | 0.00391 | 0.0101 |
| | (0.0315) | (0.0062) | (0.00999) | (0.0103) | (0.0147) |
| Lvalue_mach | 0.0526** | 0.0181* | 0.00377 | 0.00375 | 0.00202 |
| | (0.0221) | (0.0092) | (0.0114) | (0.0100) | (0.0119) |
| Risk_dummy | 0.107 | 0.126*** | 0.0648* | 0.0815** | 0.0911** |
| | (0.0889) | (0.0460) | (0.0390) | (0.0336) | (0.0409) |

| | | | | | |
|---|---|---|---|---|---|
| Lvalue_NFasset | 0.00476 | 0.0109*** | 0.0106*** | 0.0151*** | 0.0148*** |
| | (0.00680) | (0.00392) | (0.00336) | (0.00333) | (0.00411) |
| LFasset_value | 0.0415** | 0.0248*** | 0.0209*** | 0.0195*** | 0.0201** |
| | (0.0162) | (0.00942) | (0.00716) | (0.00738) | (0.00940) |
| Constant | -8.027 | 6.806 | 8.915*** | 9.644*** | 10.45*** |
| | (5.426) | (4.433) | (1.806) | (1.144) | (1.245) |
| Observations | 2,423 | 2,423 | 2,423 | 2,423 | 2,423 |

### 4.3 Income inequality decomposition

In this subsection, we study the contributions of different variables to the income inequality in China, for which we use the Shapley value approach developed by Shorrocks (1999).4 We take exponential on both sides of equation (1) in order to get the inequality decomposition for income. The decomposition is implemented using a web-based program developed by the World Institute for Development Economics Research of the United Nations University (UNUWIDER).

Table 17. Inequality decomposition results

| COMPONENTS | Whole Sample<br>Contribution to income inequality ( percent) | Urban<br>Contribution to income inequality ( percent) | Rural<br>Contribution to income inequality ( percent) |
|---|---|---|---|
| *Guanxi* index | 17.56 | 19.10 | 12.02 |
| Family Characteristic | 3.98 | 2.38 | 8.49 |
| Human Capital | 28.89 | 30.00 | 29.87 |
| Assets | 13.19 | 16.18 | 11.03 |
| Community Dummy | 36.36 | 32.35 | 38.59 |

Table 17 reports the decomposition results for the whole sample, urban and rural, respectively. The community dummies have the highest contribution among the 5 components, as it includes the regional differencies. Human capital accounts for about 30 percent of the income inequality in China. *Guanxi* index ranks as the third largest contributor to income inequalities, overtaking assets. It explains 19.10 percent of income inequality in urban China, and 12.02 percent in rural China.

We also decompose the income inequality at different quantiles. Table 18, 19 and

[4] See also Wan (2004) for a nice description of the decomposition method.

20 show the results for the whole sample, urban and rural, respectively. One can observe that the higher the quantile the greater the contribution of *Guanxi* Index for the whole sample and the urban sample. *Guanxi* even overtakes Human Capital and Assets as the second largest contributor to income inequality at 75 percent and 90 percent quantiles in urban China. *Guanxi* is not the main contributor to inequality with each income quantile in rural China.

Table 18. Inequality decomposition after quantile regression, whole sample

| COMPONENTS | q10 | q25 | q50 | q75 | q90 |
|---|---|---|---|---|---|
| *Guanxi* index | 4.73 | 9.58 | 13.95 | 12.91 | 16.28 |
| Family Characteristic | 2.94 | 4.21 | 4.49 | 4.75 | 5.82 |
| Human Capital | 28.22 | 22.83 | 22.56 | 18.09 | 15.75 |
| Assets | 15.61 | 15.74 | 14.36 | 17.44 | 17.87 |
| Community Dummy | 48.50 | 47.63 | 44.64 | 46.80 | 44.28 |

Table 19. Inequality decomposition after quantile regression, urban

| COMPONENTS | q10 | q25 | q50 | q75 | q90 |
|---|---|---|---|---|---|
| *Guanxi* index | 8.12 | 12.64 | 15.31 | 18.84 | 19.41 |
| Family Characteristic | 2.80 | 4.11 | 6.35 | 6.35 | 7.80 |
| Human Capital | 28.59 | 24.65 | 22.52 | 17.01 | 16.77 |
| Assets | 15.69 | 17.95 | 16.63 | 16.74 | 17.24 |
| Community Dummy | 44.80 | 40.65 | 39.18 | 41.07 | 38.78 |

Table 20. Inequality decomposition after quantile regression, rural

| COMPONENTS | q10 | q25 | q50 | q75 | q90 |
|---|---|---|---|---|---|
| *Guanxi* index | 2.45 | 9.01 | 6.91 | 8.01 | 7.54 |
| Family Characteristic | 4.81 | 11.89 | 5.62 | 7.93 | 7.58 |
| Human Capital | 27.58 | 25.48 | 39.30 | 15.41 | 12.29 |
| Assets | 14.49 | 17.01 | 12.34 | 21.15 | 18.65 |
| Community Dummy | 50.66 | 36.61 | 35.83 | 47.50 | 53.94 |

# 5 Conclusions

*Guanxi*, loosely translated as “connections”, is the core structure of Chinese society. It is a multidimensional concept, studied in the literature as special relationship, connection, social capital, or gift exchange economics. This paper proposes a measure of *guanxi* to capture its multiple dimensions and studies its impact on income inequality, using China household finance survey data. In line with the seminal theory

of Fei (1992), principal components analysis selects three main components to construct *Guanxi* Index: static relationships, authority and power, and means to maintain/develop connections. This paper also shows that besides wealth concentration and human capital, *guanxi* is a key factor determining income inequality, and becomes quantitatively more important than human capital for the top 25 percent of urban households.

## References


Abramson, N.R. and Ai, J.X. 1999. "Canadian companies doing business in China: Key success factors." Management International Review 39 (1): 7-35.

Benjamin, D., Brandt, L., and Giles, J. 2005. "The Evolution of Income Inequality in Rural China." *Economic Development and Cultural Change* 53(4): 769-824.

Bian, Yanjie. 1994. *Guanxi* and the allocation of urban jobs in China. China Quarterly, December, 140: 971-99.

Butterfield, Fox. 1982. "China: Alive in the Bitter Sea." New York: New York Times Books.

Coleman, James S. 1988. "Social Capital in the Creation of Human Capital." American Journal of Sociology 94: S95–121.

Farh JL, Tsui A, Xin K, Cheng BS. 1998. "The influence of relational demography and *guanxi*: the Chinese case." Organization Science 9: 471–489.

Fei, Xiaotong 1949. "Xiangtu Zhongguo." (Folk China) Beijing University Sociology Department Study Group, 1983, mimeograph. (1949 publication in China by Guan Cha She.)

Fei, Xiaotong 1992. "From the Soil: The Foundation of Chinese Society" (translated by Gary G. Hamilton and Wang Zheng). Berkeley: University of California Press.

Gold, Thomas B., Doug Guthrie, and David L. Wank, eds. 2002. "Social Connections in China: Institutions, Culture, and the Changing Nature of *Guanxi*." New York: Cambridge University Press.

Gu, F.F. Huang, K. and Tse, D.K. 2008. "When Does *Guanxi* Matter? Issues of Capitalization and Its Dark Side." Journal of Marketing, 72: 12-28.

Guo, J.X. 2005. "Human Capital, the Birth Rate and the Narrowing of the Urban-Rural Income Gap." *Social Science in China* 3: 27-37.

Guthrie, Douglas. 1998. "The Declining Significance of *Guanxi* in China's Economic Transition." China Quarterly, June, 158: 369-93.

He Qinglian. 1998. "Xiandaihua de xianjing: dangdai zhongguo de jingji shehui wenti" (Pitfalls of Modernization: Economic and Social Problems in Contemporary China). Beijing: Jinri zhongguo chubanshe.

Heshmati A. 2007. "Income Inequality in China", in Heshmati (Ed.), "Recent

Developments in the Chinese Economy", Nova Science Publishers, NY.

Hwang ER. 1987. "Face and favor: the Chinese power game. American Journal of Sociology" 92: 35–41.

Jacobs, J. B. 1979. "A preliminary model of particularistic ties in Chinese political alliances: kan-chi'ing and kuan-hsi in a rural Taiwanese township." The China Quarterly 79:237-273.

Jacobs J.B. 1980. "Local Politics in a Rural Chinese Cultural Setting: A Field Study of Mazu,Township,Taiwan." Canberra: Contemporary China Centre, Research School of Pacific Studies, Australian National University.

Khan, A.R. and Riskin, C. 1998. "Income and Inequality in China: Composition, Distribution and Growth of Household Income, 1988 to 1995." The China Quarterly 154: 221-253.

Knight, J. and Song, L. 2003. "Increasing Urban Wage Inequality in China." *Economics of Transition* 11(4): 597-619.

Knight, J. and Song, L. 1991. "The Determinants of Urban Income Inequality in China." Oxford Bulletin of Economics and Statistics 53(2): 123-154.

Knight J.B. and Yueh, L. 2008. "The Role of Social Capital in the Labor Market in China, Economics of Transition." 16(3): 384-414.

Liang, S.M. 1949. "Zhongguo wenhua yaoyi" (The Substance of Chinese Culture), Chengdu: Luming shudian.

Lin, N. 2001. "Social Capital: A Theory of Social Structure and Action." Cambridge: Cambridge University Press.

Liu, B. 1983. "People or Monster? And Other Stories and Reportage from China and Mao." (ed. by P. Link). Bloomington, Ind. Indiana University Press

Luo, Y. 1997a. "*Guanxi* and performance of foreign-invested enterprises in China: an empirical inquiry." Management International Review, 37(1): 51-71.

Luo, Y. 1997b. "Partner selection and venturing success: The case of joint ventures with firms in the People´s Republic of China." Organization Science 8 (6): 648-662.

Morduch, J. and Sicular, T. 2000. "Politics, Growth, and Inequality in Rural China : Does It Pay to Join the Party?" Journal of Public Economics 77(3): 331-356.

Morduch Jonathan and Sicular Terry 2002. “Rethingking Inequality Decomposition, with Evidence From Rural China.” The Economic Journal 112: 93-106.

North, Douglass C. 1990. “Institutions, Institutional Change and Economic Performance.” Cambridge: Cambridge University Press.

Peng, Y. 2004. “Kinship networks and entrepreneurs in China’s transitional economy.” The American Journal of Sociology,109(5): 1045–1074.

Putterman, L. 1992. “Dualism and Reform in China.” Economic Development and Cultural Change 40(3): 467-493.

Robert I. Lerman and Shlomo Y. 1985. “Income Inequality Effects by Income Source: A New Approach and Applications to the United States.” The Review of Economics and Statistics 67(1): 151-156.

Su, B.W. and Heshmati, A. 2013. “Analysis of the Determinants of Income and Income Gap between Urban and Rural China.” IZA Discussion Paper No. 7162, The Institute for the Study of Labor (IZA) in Bonn.

Tsang, Eric W.K. 1998. “Can *Guanxi* Be a Source of Sustained Competitive Advantage for Doing Business in China,” Academy of Management Executive 12 (2): 64–72.

Wan, G. 2004. “Accounting for income inequality in rural China: a regression-based approach.” Journal of Comparative Economics 32: 348–363.

Wan, G. and Zhou, Z. 2005. “Income Inequality in Rural China: Regression-Based Decomposition Using Household Data.” Review of Development Economics 9(1): 107-120.

Wei, H. and Zhao, C. 2012. “Effects of International Trade on Urban-Rural Gap income in China.” Finance & Trade Economics 1: 78-86.

Wu, X.M. and Perloff, J.M. 2005. “China’s Income Distribution, 1985-2001.” The Review of Economics and Statistics 87(4), 763-775.

Xin, K.R. and Pearce, J.L. 1996. “*Guanxi*: connections as substitutes for formal institutional support.” Academy of Management Journal, 39(6): 1641-1659.

Yang, D.T 1999. “Urban-Biased Policies and Rising Income Inequality: Emerging Pattern in China’s Reforming Economy.” Journal of Comparative Economics 19: 362-391.

Yang, D.T. and Zhou, H. 1999. “Rural-urban disparity and sectoral labour allocation in China.” Journal of Development Studies 35(3): 105-133.

Yang, M. 1994. “Gifts, favors and banquets: The art of social relationships in China.” Ithaca, NY: Cornell University.

Yeung, I. Y. M. and Tung, R. L. 1996. “Achieving business success in Confucian societies: the importance of *guanxi*.” Organisational Dynamics, 25(2): 54-65.

Zhang, Q. 2004. “Development of Financial Intermediaries and Urban-Rural Income Inequality in China.” China Journal of Finance 11: 71-79.

Zhang X.B. and Li, G. 2003. “Does *guanxi* matter for nonfarm employment?” Journal of Comparative Economics, 31(2): 315-331.